GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 08:52:43 ; Search time 374.18 Seconds
(without alignments)
10074.554 Million cell updates/sec

Title: US-08-852-495C-1_COPY_140000_169000
Perfect score: 29001
Sequence: 1 CCCTCCAATCCCATATGCAC.....TGTTATCACAGAAAGTTACC 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	76.8	0.3	105	4	US-08-481-658B-65	Sequence 65, Appl
c 2	76.8	0.3	105	4	US-08-477-504A-65	Sequence 65, Appl
c 3	76.8	0.3	105	4	US-08-486-756A-65	Sequence 65, Appl
c 4	76.8	0.3	105	4	US-08-485-862B-65	Sequence 65, Appl
c 5	76.8	0.3	105	5	US-08-787-739-65	Sequence 65, Appl
6	71.2	0.2	105	4	US-08-481-658B-65	Sequence 65, Appl
7	71.2	0.2	105	4	US-08-477-504A-65	Sequence 65, Appl
8	71.2	0.2	105	4	US-08-486-756A-65	Sequence 65, Appl
9	71.2	0.2	105	4	US-08-485-862B-65	Sequence 65, Appl
10	71.2	0.2	105	5	US-08-787-739-65	Sequence 65, Appl
c 11	65.4	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
c 12	65.4	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
c 13	65.4	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
c 14	65.4	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
15	60.4	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
16	60.4	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
17	60.4	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
18	60.4	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
19	57	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
20	57	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
21	57	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
22	57	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
c 23	56	0.2	85	3	US-08-454-557C-92	Sequence 92, Appl
c 24	56	0.2	85	4	US-08-340-426D-92	Sequence 92, Appl
c 25	56	0.2	85	4	US-08-450-673C-92	Sequence 92, Appl
c 26	56	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
c 27	55.6	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
c 28	55.6	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
c 29	55.6	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
c 30	55.6	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
31	55.2	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
32	55.2	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
33	55.2	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
34	55.2	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
35	55	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
36	55	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
37	55	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
38	55	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
c 39	53.2	0.2	83	4	US-08-481-658B-66	Sequence 66, Appl
c 40	53.2	0.2	83	4	US-08-477-504A-66	Sequence 66, Appl
c 41	53.2	0.2	83	4	US-08-486-756A-66	Sequence 66, Appl
c 42	53.2	0.2	83	4	US-08-485-862B-66	Sequence 66, Appl
c 43	53.2	0.2	83	5	US-08-787-739-66	Sequence 66, Appl
c 44	50	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
c 45	50	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 4e-07;

```
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  24042 ATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGATCAAGACCATC 24101
          |||||||||||||||||||| ||||  |||| ||||| |||||||||| |  ||| || |
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  24102 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAATACAAAAAA 24145
          |||||||| |||||||||||| ||||||||||||     |||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   2
US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.3%;  Score 76.8;  DB 4;  Length 105;
  Best Local Similarity   83.7%;  Pred. No. 4e-07;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  24042 ATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGATCAAGACCATC 24101
          |||||||||||||||||||| ||||  |||| ||||| |||||||||| |  ||| || |
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  24102 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAATACAAAAAA 24145
          |||||||| |||||||||||| ||||||||||||     |||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   3
US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.3%;  Score 76.8;  DB 4;  Length 105;
  Best Local Similarity   83.7%;  Pred. No. 4e-07;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  24042 ATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGATCAAGACCATC 24101
          |||||||||||||||||||| ||||  |||| ||||| |||||||||| |  ||| || |
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  24102 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAATACAAAAAA 24145
          |||||||| |||||||||||| ||||||||||||     |||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   4
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
```

```
;       CITY:  Tiburon
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94920
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,862B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;       APPLICATION NUMBER:  US 08/260,190
;       FILING DATE:  15-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.3D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.3%;  Score 76.8;  DB 4;  Length 105;
  Best Local Similarity   83.7%;  Pred. No. 4e-07;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  24042 ATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGATCAAGACCATC 24101
          ||||||||||||||||||||| ||||  |||| ||||| |||||||||| |  ||| || |
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  24102 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAATACAAAAAA 24145
          ||||||| |||||||||||| ||||||||||||     |||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   5
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-787-739-65


  Query Match             0.3%;  Score 76.8;  DB 5;  Length 105;
  Best Local Similarity   83.7%;  Pred. No. 4e-07;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  24042 ATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGATCAAGACCATC 24101
          ||||||||||||||||||||| ||||  |||| ||||| |||||||||| |  ||| || |
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  24102 CTGGCCAACATGGTGAAACCCCGTCTCTACTAAAATACAAAAAA 24145
          ||||||| |||||||||||| ||||||||||||     |||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   6
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match             0.2%;  Score 71.2;  DB 4;  Length 105;
  Best Local Similarity   85.8%;  Pred. No. 5.4e-06;
  Matches   91;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy  16710 TTTTTGTATTTTTAGTAGAGATAGGGTTTCACAATGCTGGCCAGGCTGGTCTCAAACTCC 16769
          |||||  || ||||||||||| |||||||||| ||  |||||||||| ||||||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  16770 TGCCCTCAAGTGATCCTCCTGCCTCGGCCTCCCAATGTGCTGGGAT 16815
          || |||   ||||||| || ||||||||||||||| ||||||||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   7
US-08-477-504A-65
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.2%;  Score 71.2;  DB 4;  Length 105;
  Best Local Similarity   85.8%;  Pred. No. 5.4e-06;
  Matches   91;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy  16710 TTTTTGTATTTTTAGTAGAGATAGGGTTTCACAATGCTGGCCAGGCTGGTCTCAAACTCC 16769
          |||||  || ||||||||||| |||||||||| ||  |||||||||| ||||||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  16770 TGCCCTCAAGTGATCCTCCTGCCTCGGCCTCCCAATGTGCTGGGAT 16815
          || |||   ||||||| || ||||||||||||||| ||||||||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   8
US-08-486-756A-65
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.2%;  Score 71.2;  DB 4;  Length 105;
  Best Local Similarity   85.8%;  Pred. No. 5.4e-06;
  Matches   91;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy  16710 TTTTTGTATTTTTAGTAGAGATAGGGTTTCACAATGCTGGCCAGGCTGGTCTCAAACTCC 16769
          |||||  || ||||||||||| |||||||||| ||  |||||||||| ||||||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  16770 TGCCCTCAAGTGATCCTCCTGCCTCGGCCTCCCAATGTGCTGGGAT 16815
          || |||   ||||||| || ||||||||||||||| ||||||||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  9
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,862B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.2%;  Score 71.2;  DB 4;  Length 105;
  Best Local Similarity   85.8%;  Pred. No. 5.4e-06;
  Matches   91;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy  16710 TTTTTGTATTTTTAGTAGAGATAGGGTTTCACAATGCTGGCCAGGCTGGTCTCAAACTCC 16769
          |||||  || ||||||||||| |||||||||| ||  |||||||||| ||||||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  16770 TGCCCTCAAGTGATCCTCCTGCCTCGGCCTCCCAATGTGCTGGGAT 16815
          || |||   ||||||| || ||||||||||||||| ||||||||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  10
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-787-739-65
```

```
  Query Match             0.2%;  Score 71.2;  DB 5;  Length 105;
  Best Local Similarity   85.8%;  Pred. No. 5.4e-06;
  Matches   91;  Conservative    0;  Mismatches   13;  Indels    2;  Gaps    1;

Qy  16710 TTTTTGTATTTTTAGTAGAGATAGGGTTTCACAATGCTGGCCAGGCTGGTCTCAAACTCC 16769
          |||||  || |||||||||| |||||||||| ||  |||||||||| ||||||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  16770 TGCCCTCAAGTGATCCTCCTGCCTCGGCCTCCCAATGTGCTGGGAT 16815
          || |||   ||||||| || ||||||||||||| ||||||||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  11
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-91


  Query Match             0.2%;  Score 65.4;  DB 3;  Length 84;
  Best Local Similarity   86.7%;  Pred. No. 7.4e-05;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  24032 CACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGAT 24091
          ||||| |||||||||||||||||||||||||| ||| | |||||||||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  24092 CAAGACCATCCTGGCCAACATGG 24114
          | | |||| ||||   |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  12
US-08-340-426D-91/c
; Sequence 91, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-91


  Query Match             0.2%;  Score 65.4;  DB 4;  Length 84;
  Best Local Similarity   86.7%;  Pred. No. 7.4e-05;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  24032 CACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGAT 24091
          ||||| |||||||||||||||||||||||||| ||| | |||||||||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  24092 CAAGACCATCCTGGCCAACATGG 24114
          | | |||| ||||   |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  13
US-08-450-673C-91/c
; Sequence 91, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-91


  Query Match             0.2%;  Score 65.4;  DB 4;  Length 84;
  Best Local Similarity   86.7%;  Pred. No. 7.4e-05;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  24032 CACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGAT 24091
          ||||| |||||||||||||||||||||||||||| ||| | ||||||||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  24092 CAAGACCATCCTGGCCAACATGG 24114
          | | |||| ||||   |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  14
PCT-US95-17111A-91/c
; Sequence 91, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-91


  Query Match             0.2%;  Score 65.4;  DB 6;  Length 84;
  Best Local Similarity   86.7%;  Pred. No. 7.4e-05;
  Matches   72;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  24032 CACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGAATCACGAGGTCAGGAGAT 24091
          ||||| |||||||||||||||||||||||||||| ||| | ||||||||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  24092 CAAGACCATCCTGGCCAACATGG 24114
          | | |||| ||||   |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  15
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-70


  Query Match             0.2%;  Score 60.4;  DB 3;  Length 78;
  Best Local Similarity   85.9%;  Pred. No. 0.00074;
  Matches   67;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   9486 ACCATGCCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGG 9545
          || | |||| |||||| |||||||||||||||||  ||||||| |||||||   ||||
Db      1 ACAACGCCCAGCTAATATTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTCATCAGG 60

Qy   9546 CTGGTGTTGAACTCCTGA 9563
          ||||||| ||||||||||
Db     61 CTGGTGTCGAACTCCTGA 78
```

Search completed: June 16, 2000, 20:15:34
Job time: 199264 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 20:09:08 ; Search time 29137.4 Seconds
(without alignments)
-968.237 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTTAGAGGAAAAAAAGCA.....TAGATAAAACGTTGTTCCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	95.2	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
	2	92	0.3	108	10	HSLDLRN2	X05250 Human LDL-r
	3	85	0.3	107	9	HUMALCE162	M87924 Human carci
	4	84.6	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
c	5	84.6	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
c	6	83	0.3	108	10	HSLDLRD1	X05249 Human LDL-r
	7	83	0.3	108	10	HSLDLRD2	X05251 Human LDL-r
	8	82.2	0.3	103	9	HUMALCE221	M87896 Human carci
	9	81	0.3	108	11	HSU67803	U67803 Human small
	10	81	0.3	108	11	HSU67808	U67808 Human small
c	11	79.2	0.3	103	13	HS8IC8R	X57789 Human seque
c	12	78	0.3	107	9	HUMALCE162	M87924 Human carci
c	13	77.6	0.3	103	9	HUMALCE221	M87896 Human carci
	14	77.2	0.3	110	11	HSU67807	U67807 Human small
	15	76.6	0.3	108	10	HSLDLI12	X05248 Human LDL-r
c	16	76.6	0.3	110	11	HSU67807	U67807 Human small
	17	75.2	0.3	97	9	HUMLDLRA2	M14180 Human low d
	18	74.4	0.3	103	13	HS8IC8R	X57789 Human seque
	19	74.4	0.3	104	9	HUMALCE272	M87899 Human carci
	20	74.6	0.3	108	11	HSU67804	U67804 Human small
	21	73.4	0.3	91	13	HUMUT8164A	L30244 Human STS U
c	22	72.4	0.2	91	13	HUMUT8164A	L30244 Human STS U
c	23	72.4	0.2	108	10	HSLDLI12	X05248 Human LDL-r
	24	72.4	0.2	110	9	HUMALCE43	M87900 Human carci
c	25	72	0.2	97	9	HUMLDLRA2	M14180 Human low d
c	26	72.2	0.2	108	9	HUMD1D03M5	D16965 Human HepG2
	27	71.4	0.2	107	11	HSU67806	U67806 Human small
	28	71.4	0.2	108	13	G43535	G43535 WIAF-2393-S
c	29	70.8	0.2	90	9	HUMLDLRFL	K03555 Human low d
c	30	70.8	0.2	100	13	HUMUT931A	L31299 Human STS U
	31	70.2	0.2	97	9	HUMLDLRDJ	M14179 Human famil
c	32	70.4	0.2	106	13	G32743	G32743 A009P31 Hum
c	33	69.4	0.2	108	13	G43535	G43535 WIAF-2393-S
c	34	68.6	0.2	84	5	AR051521	AR051521 Sequence
	35	68.4	0.2	95	13	HUMUT8002B	L30176 Human STS U
	36	68.2	0.2	99	13	HUMUT7692A	L30306 Human STS U
	37	67.8	0.2	108	9	HUMD1D03M5	D16965 Human HepG2
	38	67.4	0.2	79	10	S73203	S73203 ALL-1 (tand
c	39	67.4	0.2	102	13	G32906	G32906 A009W09 Hum
	40	66.8	0.2	95	10	HSSTHPKIB	X66361 H.sapiens m
c	41	67	0.2	97	9	HUMLDLRDJ	M14179 Human famil
	42	66	0.2	97	9	HUMLDLRA1	M14178 Human low d
	43	65.8	0.2	106	13	G32743	G32743 A009P31 Hum
c	44	64.8	0.2	79	10	S73203	S73203 ALL-1 (tand
c	45	65	0.2	99	13	HUMUT7692A	L30306 Human STS U

ALIGNMENTS

```
RESULT   1
HSLDLRN2/c
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.3%;  Score 95.2;  DB 10;  Length 108;
  Best Local Similarity   92.6%;  Pred. No. 3e-06;
  Matches  100;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy  22680 CTTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCCCATGCCTCAGCCTCCCG 22739
          || |||||||||||||||| |||||||||||||||||||| | ||||||||||||||
Db    108 CTCGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCG 49

Qy  22740 AGTAGCTGGGATTACAGGCACATACCACCATGACTGGCTAATTTTTGT 22787
          |||||||||||||||||||| | |||||| | ||||||||||||||
Db     48 AGTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   2
HSLDLRN2
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.3%;  Score 92;  DB 10;  Length 108;
  Best Local Similarity   90.7%;  Pred. No. 1e-05;
  Matches   98;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16843 ACAAAAATTAGCCAGGCCTGGTGGCATCTGCCTGTAGTCCCAGCTACTCGGGACACTGAG 16902
          |||||||||||||||| ||||||||  |||||||| |||||||||||||||  |||||
Db      1 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 60

Qy  16903 GCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGAG 16950
          |||||||||  |||||||||||||||||||| ||||||||||  |||
Db     61 GCAGGAGAATTGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAG 108


RESULT   3
HUMALCE162
LOCUS       HUMALCE162    107 bp ss-RNA           PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION   M87924
VERSION     M87924.1  GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       28 a     30 c     35 g     14 t
ORIGIN


  Query Match             0.3%;  Score 85;  DB 9;  Length 107;
  Best Local Similarity   90.1%;  Pred. No. 0.00014;
  Matches   91;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16902 GGCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGAGATCGCGCCACT 16961
          ||||| |||||  | ||||||| |||||| ||| ||||||||||| ||||||||||||||
Db      5 GGCAGAAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACT 64

Qy  16962 GCATTCCAGCCTGGGAGACAGAGCGAGACTCCGTCTCAAAA 17002
          ||| |||||||||||| |||||||||||||||||||||||||
Db     65 GCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAA 105


RESULT   4
HSLDLRD1
LOCUS       HSLDLRD1      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION   X05249
VERSION     X05249.1  GI:34335
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
```

```
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05248 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     misc_feature    1. .108
                     /note="deletion junction region intron 12/ intron 15"
BASE COUNT       20 a     40 c     20 g     28 t
ORIGIN


  Query Match             0.3%;  Score 84.6;  DB 10;  Length 108;
  Best Local Similarity   86.9%;  Pred. No. 0.00016;
  Matches   93;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  22681 TTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCCCATGCCTCAGCCTCCCGA 22740
          | | |||||  ||||||| ||||||||||||||| | |||  | ||||||||||||||||
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 61

Qy  22741 GTAGCTGGGATTACAGGCACATACCACCATGACTGGCTAATTTTTGT 22787
          ||||||||||||||||||| | |||||| | ||||||||||||||||
Db     62 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 108


RESULT   5
HSLDLRD2/c
LOCUS       HSLDLRD2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION   X05251
VERSION     X05251.1  GI:34336
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     20 c     40 g     20 t
ORIGIN


  Query Match             0.3%;  Score 84.6;  DB 10;  Length 108;
  Best Local Similarity   86.9%;  Pred. No. 0.00016;
  Matches   93;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  22681 TTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTCCCATGCCTCAGCCTCCCGA 22740
          | | |||||  ||||||| ||||||||||||||| | |||  | ||||||||||||||||
Db    107 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 48

Qy  22741 GTAGCTGGGATTACAGGCACATACCACCATGACTGGCTAATTTTTGT 22787
          ||||||||||||||||||| | |||||| | ||||||||||||||||
Db     47 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   6
HSLDLRD1/c
LOCUS       HSLDLRD1      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION   X05249
VERSION     X05249.1  GI:34335
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05248 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     misc_feature    1. .108
                     /note="deletion junction region intron 12/ intron 15"
BASE COUNT       20 a     40 c     20 g     28 t
ORIGIN


  Query Match             0.3%;  Score 83;  DB 10;  Length 108;
  Best Local Similarity   86.0%;  Pred. No. 0.00029;
  Matches   92;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  16843 ACAAAAATTAGCCAGGCCTGGTGGCATCTGCCTGTAGTCCCAGCTACTCGGGACACTGAG 16902
          |||||||||||||||| |||||||| |||||||| |||||||||||||||  |||||
Db    108 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 49

Qy  16903 GCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGA 16949
          |||||| |||   |||||||||||||||||| |||  |||||  ||
Db     48 GCAGGAAAATGGTTTGAACCCAGGAGGCAGAGGTTGTGGTGAGGCGA 2


RESULT   7
HSLDLRD2
LOCUS       HSLDLRD2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION   X05251
VERSION     X05251.1  GI:34336
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
```

```
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     20 c     40 g     20 t
ORIGIN


  Query Match             0.3%;  Score 83;  DB 10;  Length 108;
  Best Local Similarity   86.0%;  Pred. No. 0.00029;
  Matches   92;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  16843 ACAAAAATTAGCCAGGCCTGGTGGCATCTGCCTGTAGTCCCAGCTACTCGGGACACTGAG 16902
          ||||||||||||||||| ||||||||  |||||||| ||||||||||||||||  |||||
Db      1 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 60

Qy  16903 GCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGA 16949
          |||||| |||   ||||||||||||||||||| |||  |||||  ||
Db     61 GCAGGAAAATGGTTTGAACCCAGGAGGCAGAGGTTGTGGTGAGGCGA 107


RESULT   8
HUMALCE221
LOCUS       HUMALCE221    103 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION   M87896
VERSION     M87896.1  GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       25 a     27 c     33 g     18 t
ORIGIN


  Query Match             0.3%;  Score 82.2;  DB 9;  Length 103;
  Best Local Similarity   87.4%;  Pred. No. 0.0004;
  Matches   90;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  16868 ATCTGCCTGTAGTCCCAGCTACTCGGGACACTGAGGCAGGAGAATCACTTGAACCCAGGA 16927
          |||||||||||| |||||||||| |||||  || |||||||||||||| ||||||||| |||
Db      1 ATCTGCCTGTAATCCCAGCTACACGGGAAGCTAAGGCAGGAGAATCGCTTGAACCCGGGA 60

Qy  16928 GGCAGAGATTGCAGTGAGCTGAGATCGCGCCACTGCATTCCAG 16970
          ||| ||| |||||||||||| ||||||| |||| |||| |||||
Db     61 GGCGGAGGTTGCAGTGAGCCGAGATCGTGCCATTGCACTCCAG 103


RESULT   9
HSU67803
LOCUS       HSU67803      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67803
VERSION     U67803.1  GI:2289917
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu2"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       23 a     39 c     30 g     16 t
ORIGIN


  Query Match             0.3%;  Score 81;  DB 11;  Length 108;
  Best Local Similarity   89.7%;  Pred. No. 0.00062;
  Matches   87;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16737 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGA 16796
          |||||||||||||||||||||||||||| |||||| | |||||||||||||||| | ||
Db      1 GCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCGA 60

Qy  16797 GACCAGCCTGACCAACATGGTGAAACCCTGTCTCTAC 16833
          ||||| |||| | |||| |||||||||| ||||||||
Db     61 GACCATCCTGGCTAACAAGGTGAAACCCCGTCTCTAC 97


RESULT  10
HSU67808
LOCUS       HSU67808      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67808
VERSION     U67808.1  GI:2289922
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
```

```
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu7"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       22 a     37 c     28 g     21 t
ORIGIN


  Query Match             0.3%;  Score 81;  DB 11;  Length 108;
  Best Local Similarity   89.7%;  Pred. No. 0.00062;
  Matches   87;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16737 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGA 16796
          ||||||||||||||||||||||||||||||| |||   |||||| ||||||||||||||||
Db      1 GCCTGTAATCCCAGCACTTTGGGAGGCCAAGTCGGGTGGATCACAAGGTCAGGAGTTTGA 60

Qy  16797 GACCAGCCTGACCAACATGGTGAAACCCTGTCTCTAC 16833
          ||||||||| |||||||||||||| | |||| | |
Db     61 GACCAGCCTGGCCAACATGGTGAAACTCCGTCTTTCC 97


RESULT  11
HS8IC8R/c
LOCUS       HS8IC8R       103 bp    DNA             STS       05-SEP-1991
DEFINITION  Human sequence tagged site 8IC8R DNA from 19q13.
ACCESSION   X57789
VERSION     X57789.1  GI:23938
KEYWORDS    STS; myotonic dystrophy.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Aldridge,F.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
            Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE   2  (bases 1 to 103)
  AUTHORS   Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
            Davies,J., Johnson,K. and Markham,A.F.
  TITLE     Two sequence-tagged sites defining the ends of a 380 kb YAC clone
            from 19q13
  JOURNAL   Nucleic Acids Res. 19 (17), 4787 (1991)
  MEDLINE   91367697
COMMENT     See also X57788 for STS 8IC8L.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19q13"
                     /germline
                     /clone_lib="YAC library: ICI"
                     /clone="8IC8"
BASE COUNT       29 a     28 c     23 g     22 t      1 others
ORIGIN


  Query Match             0.3%;  Score 79.2;  DB 13;  Length 103;
  Best Local Similarity   86.1%;  Pred. No. 0.0012;
  Matches   87;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  27423 TACCTGTAATCCCAGCACTTTGGGAGACTGAGGTGGGTGGATCACTTGAGGTCAGGAGTT 27482
          | ||| ||||  | ||||||||||||   ||||||||||||||||||| ||||||||||||

Db    102 TGCCTATAATTTCNGCACTTTGGGAGGTCGAGGTGGGTGGATCACTTAAGGTCAGGAGTT 43

Qy  27483 CAAAACCAGCCTGGCCAACATGGTGAAAACCCATCTCTACT 27523
          |  |||||||||||||||||||||||| || |||||||||
Db     42 CTTGACCAGCCTGGCCAACATGGTGAAACCCTATCTCTACT 2


RESULT  12
HUMALCE162/c
LOCUS       HUMALCE162    107 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION   M87924
VERSION     M87924.1  GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       28 a     30 c     35 g     14 t
ORIGIN


  Query Match             0.3%;  Score 78;  DB 9;  Length 107;
  Best Local Similarity   85.3%;  Pred. No. 0.0019;
  Matches   87;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  18560 TTTTTGAGACAGAGTCTGACTCTGTTGCCCAGGCTGGAGTGCAGTGGCACCATCTAGGCT 18619
          |||||||||| ||||||  |||||| |||||||||||||||||||||| | |||| ||||
Db    106 TTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCT 47

Qy  18620 CACTGCAGCCTCTACCTCTCGGGTTCAAGCGATTCTCCTGCC 18661
          |||||||  |||  |||| |||||||| || ||||| |||||
Db     46 CACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTTCTGCC 5


RESULT  13
HUMALCE221/c
LOCUS       HUMALCE221    103 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION   M87896
VERSION     M87896.1  GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       25 a     27 c     33 g     18 t
```

```
ORIGIN


  Query Match             0.3%;  Score 77.6;  DB 9;  Length 103;
  Best Local Similarity   86.0%;  Pred. No. 0.0022;
  Matches   86;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  10572 CTGGAGTGCAGTGGCACGATCTCAGCTCACTGAAACCTCCAATTTCTGAGTTCAAGCGAT 10631
          |||||||||| |||||||||||| |||||||| |||||||   | | | ||||||||||||
Db    103 CTGGAGTGCAATGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAT 44

Qy  10632 TCTCGTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCA 10671
          |||| ||||| ||| |||| ||||||||||||||||||
Db     43 TCTCCTGCCTTAGCTTCCCGTGTAGCTGGGATTACAGGCA 4


RESULT  14
HSU67807
LOCUS       HSU67807      110 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67807
VERSION     U67807.1  GI:2289921
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 110)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu6"
     repeat_region   1. .110
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       26 a     39 c     24 g     21 t
ORIGIN


  Query Match             0.3%;  Score 77.2;  DB 11;  Length 110;
  Best Local Similarity   86.7%;  Pred. No. 0.0026;
  Matches   85;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  27425 CCTGTAATCCCAGCACTTTGGGAGACTGAGGTGGGTGGATCACTTGAGGTCAGGAGTTCA 27484
          |||||||||||||||||| ||| |  | ||||||||||||||||||  ||||||||||
Db      2 CCTGTAATCCCAGCACTTTGAGAGGCCAAAGTGGGTGGATCACTTGAGCCCAGGAGTTCA 61

Qy  27485 AAACCAGCCTGGCCAACATGGTGAAAACCCATCTCTAC 27522
          | ||||||||| | |||||||||||| ||||||| | |
Db     62 AGACCAGCCTGTCAAACATGGTGAAACCCCATCTTTCC 99


RESULT  15
HSLDLI12
LOCUS       HSLDLI12      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low
            density lipoprotein.
ACCESSION   X05248
VERSION     X05248.1  GI:34334
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor;
            repetitive sequence.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     see X05249 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    complement(<1. .65)
                     /note="Alu repeat"
     intron          1. .108
                     /note="intron XII fragment"
BASE COUNT       21 a     38 c     20 g     29 t
ORIGIN


  Query Match             0.3%;  Score 76.6;  DB 10;  Length 108;
  Best Local Similarity   82.2%;  Pred. No. 0.0032;
  Matches   88;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy  15320 TTGGTTCACTGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTTCTGCCTCAGTCTCCTGA 15379
          | |  ||||  |||||||||||||||||||||||| | ||| | ||||||||| ||||| |
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCTTA 61

Qy  15380 GGAGCTGGGACTACAGACATGTGCCACTACACCCAGCTAATTTTGTA 15426
          | |||||||| ||||  |||||||||| || ||| ||| ||||||||
Db     62 GTAGCTGGGATTACAAGCATGTGCCACCACGCCCGGCTGATTTTGTA 108



Search completed: June 17, 2000, 10:40:12
Job time: 252096 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 21:16:14 ; Search time 939.94 Seconds
(without alignments)
7719.451 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTTAGAGGAAAAAAAGCA.....TAGATAAAACGTTGTTCCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	0.2	108	1	X12095	Human biallelic po
c 2	70	0.2	108	1	X12095	Human biallelic po
c 3	67.6	0.2	108	1	T25009	Human gene signatu
4	65.6	0.2	100	1	T24892	Human gene signatu
c 5	63.4	0.2	88	1	V39744	Microsatellite ana
c 6	62.4	0.2	100	1	T24892	Human gene signatu
7	62.6	0.2	103	1	T26213	Human gene signatu
8	62.4	0.2	108	1	T25009	Human gene signatu
c 9	61.6	0.2	87	1	T21566	Human gene signatu
10	61.8	0.2	103	1	T20927	Human gene signatu
c 11	61.4	0.2	108	1	T26828	Human gene signatu
12	61	0.2	91	1	T25854	Human gene signatu
13	61	0.2	108	1	T26828	Human gene signatu
c 14	60.6	0.2	93	1	T22572	Human gene signatu
c 15	60.4	0.2	103	1	T26213	Human gene signatu
16	58.6	0.2	100	1	X12087	Human biallelic po
17	58.6	0.2	100	1	X12085	Human biallelic po
18	58	0.2	87	1	T21566	Human gene signatu
19	57.4	0.2	93	1	T24259	Human gene signatu
20	57.4	0.2	100	1	X12086	Human biallelic po
21	56.6	0.2	92	1	T66081	(dC-dA)n.(dG-dT)n
22	56.4	0.2	93	1	T22572	Human gene signatu
c 23	55.8	0.2	93	1	T24259	Human gene signatu
c 24	55.4	0.2	81	1	T24093	Human gene signatu
c 25	55.6	0.2	95	1	T23131	Human gene signatu
c 26	55	0.2	99	1	T24420	Human gene signatu
c 27	54.6	0.2	91	1	T25854	Human gene signatu
28	54	0.2	60	1	T65762	Repeat sequence fr
29	53.8	0.2	99	1	T20931	Human gene signatu
c 30	54	0.2	100	1	X12087	Human biallelic po
c 31	54	0.2	100	1	X12085	Human biallelic po
c 32	54	0.2	100	1	X12086	Human biallelic po
c 33	53.6	0.2	69	1	Q29016	Probe to internal
34	53.6	0.2	91	1	T65740	Repeat sequence fr
c 35	53.4	0.2	99	1	T23728	Human gene signatu
c 36	53	0.2	97	1	T26728	Human gene signatu
37	52	0.2	69	1	Q29016	Probe to internal
c 38	51.6	0.2	103	1	T20927	Human gene signatu
39	51	0.2	84	1	T25848	Human gene signatu
40	51.2	0.2	93	1	T25688	Human gene signatu
41	50.8	0.2	97	1	T26728	Human gene signatu
42	50.6	0.2	110	1	T26288	Human gene signatu
c 43	50.8	0.2	110	1	T26288	Human gene signatu
c 44	50.2	0.2	64	1	Q33534	Microsatellite seq
45	50	0.2	75	1	T22841	Human gene signatu

ALIGNMENTS

RESULT 1
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.2%; Score 72; DB 1; Length 108;
Best Local Similarity 79.2%; Pred. No. 0.011;
Matches 84; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

```
Qy   8485 TTTTTTTAATAGAGATGGGGTTTTGCCATGTTAACCAGGCAGGTCTTGAACTCCTGACCT 8544
          | |||||  |||||||| |||||| | :|||| |||||  ||||| ||||||||||| |
Db      3 TCTTTTTTGTAGAGATGAGGTTTTCCTRTGTTGGCCAGGATGGTCTCGAACTCCTGACTT 62

Qy   8545 CAAGTGATCTGCCCACCTTGGCCCCTCCATAGTGCTGGGATTACAG 8590
          |||||||||| | |  |||||||| | | | ||||||||||||| ||
Db     63 CAAGTGATCCGTCTGCCTTGGCCTCCCAAAAGTGCTGGGATTATAG 108
```

```
RESULT   2
X12095/c
ID   X12095 standard; DNA; 108 BP.
AC   X12095;
DT   30-MAR-1999  (first entry)
DE   Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW   Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW   detection; phenotypic typing; characteristic; infection; hereditary;
KW   autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW   treatment; marker; ss.
OS   Homo sapiens.
PN   WO9820165-A2.
PD   14-MAY-1998.
PF   05-NOV-1997; U20313.
PR   06-NOV-1996; US-030455.
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI   Hudson T, Lander ES, Wang D;
DR   WPI; 98-286974/25.
PT   New isolated nucleic acid segments from the human genome - used for
PT   determining polymorphic forms for use in e.g. forensics, paternity
PT   testing or phenotypic typing for disease
PS   Claim 1; Page 219; 310pp; English.
CC   X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC   markers which have been isolated using the primers represented in
CC   X09121-X10268. The base occupying the polymorphic site is indicated by
CC   the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC   methods for determining polymorphic forms in an individual for use in
CC   e.g. forensics, paternity testing or for phenotypic typing for diseases
CC   such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC   muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC   hypercholesterolemia, polycystic kidney disease, hereditary
CC   spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC   haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC   syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC   autoimmune diseases, inflammation, cancer, diseases of the nervous
CC   system, infection by pathogenic microorganisms, and characteristics such
CC   as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC   endurance, fertility, and susceptibility or receptivity to particular
CC   drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC   segments can also be used to produce medicaments for the treatment or
CC   prophylaxis of such diseases.
SQ   Sequence   108 BP;    19 A;    23 C;    28 G;    37 T;


  Query Match             0.2%;  Score 70;  DB 1;  Length 108;
  Best Local Similarity   83.3%;  Pred. No. 0.022;
  Matches   90;  Conservative    1;  Mismatches   16;  Indels    1;  Gaps    1;

Qy   1140 CTGTAATCCTAGCAC-TTTGGGAAGCCGAGGCGGGCAGATCACCTGAGGTCAGGAGTTTG 1198
            || |||||| ||||| ||||||| ||| |||| | | |||||| ||| ||||||||||  |
Db    108 CTATAATCCCAGCACTTTTGGGAGGCCAAGGCAGACGGATCACTTGAAGTCAGGAGTTCG 49

Qy   1199 AGACCAGCCTGACCAACATGGCAAAACCTCATCTCTACTAAAAATACA 1246
           |||||| |||| ||||||: | ||||||||||||||| ||||| |||
Db     48 AGACCATCCTGGCCAACAYAGGAAAACCTCATCTCTACAAAAAAGACA 1


RESULT   3
T25009/c
ID   T25009 standard; cDNA to mRNA; 108 BP.
AC   T25009;
DT   07-NOV-1996  (first entry)
DE   Human gene signature HUMGS07131.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1748; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    34 A;    31 C;    26 G;    15 T;


  Query Match             0.2%;  Score 67.6;  DB 1;  Length 108;
  Best Local Similarity   75.9%;  Pred. No. 0.05;
  Matches   82;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy   19044 TTTTTTTTTTTTTTTTTTTTACGCAGAGTCCTACTCTGTCACCCAGGCTGGAGTGCAGCG 19103
            |||  | || || || ||||   ||| ||| | |||||||| |||||||| | ||| |
Db     108 TTTGNTGTTGTTGTTGTTTTCAACAGGGTCTTGCTCTGTCACTCAGGCTGGAATNCAGTG 49

Qy   19104 GTGTGAACATGGCTCACTGCAGCCTCAACCTACTGAGCTCAAGCAATC 19151
            | |||| ||||||||||||||||||   |||  || ||||| || |||
Db      48 GCGTGACCATGGCTCACTGCAGCCTTGGCCTCATGGGCTCAGGCGATC 1


RESULT   4
T24892
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
```

```
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;


  Query Match             0.2%;  Score 65.6;  DB 1;  Length 100;
  Best Local Similarity   77.8%;  Pred. No. 0.097;
  Matches   77;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy  27600 ATCACTTGAACTCAGGAGGCAGAGGTTGTAGTGAGCTGAGATCGCACCACTGCACTCCAG 27659
          ||| |||||||   |||||||||  ||| | | ||||||||| |||||  |||||||| |
Db      2 ATCGCTTGAACCTGGGAGGCAGAATTTGCAATNAGCTGAGATTGCACCNTTGCACTCCNG 61

Qy  27660 CCTGGGTGACAGAGCAAGACTCCATTTAAAAAAAAAATA 27698
          ||||||||||||||  | ||||  ||| ||| ||| | |
Db     62 CCTGGGTGACAGAGTGACACTCTGTTTGAAACAAACAAA 100


RESULT   5
V39744/c
ID   V39744 standard; cDNA; 88 BP.
AC   V39744;
DT   28-SEP-1998  (first entry)
DE   Microsatellite analysis antisense strand SEQ ID NO:262.
KW   Mass spectrometry; diagnosis; detection; biological sample; infection;
KW   genetic disease; chromosomal abnormality; identification; heredity;
KW   pathogenic organism; telomerase activity; oncogene mutation;
KW   cancer-specific sequence; primer; ss.
OS   Synthetic.
PN   WO9820166-A2.
PD   14-MAY-1998.
PF   06-NOV-1997; U20444.
PR   08-OCT-1997; US-947801.
PR   06-NOV-1996; US-744481.
PR   06-NOV-1996; US-744590.
PR   06-NOV-1996; US-746036.
PR   06-NOV-1996; US-746055.
PR   23-JAN-1997; US-786988.
PR   23-JAN-1997; US-787639.
PR   19-SEP-1997; US-933792.
PA   (SEQU-) SEQUENOM INC.
PI   Braun A, Damhoffer-Demar B, Fu D, Higgins GS, Jurinke C,
PI   Koster H, Little DP, Lough DM, Siegert CW, Tang K,
PI   Van Den Boom D, Xiang G;
DR   WPI; 98-286975/25.
PT   Sequencing nucleic acid by mass spectrometric analysis  - for
PT   detecting nucleic acids, telomerase activity, oncogene mutations, or
PT   cancer-specific sequences, for diagnosis of disease
PS   Example 11; Page 318; 478pp; English.
CC   A process has been developed for determining the sequence of a target
CC   nucleic acid. The process comprises: (i) generating at least two
CC   fragments (F) from the target nucleic acid; and (ii) analysing F by
CC   mass spectrometry (MS). The sequences in V39483 to V39592 are
CC   specifically claimed primers for use in the mass spectrometric analysis
CC   of the above process. The process is used to detect genetic diseases
CC   (e.g. haemophilia, thalassemia, Duchenne muscular dystrophy, Alzheimer's
CC   disease, cystic fibrosis and many others) or chromosomal abnormalities
CC   (or predisposition); infections and cancers; also for establishing
CC   identity and heredity. Particular applications are diagnosis of
CC   neuroblastoma, detecting telomerase, determining family relationships
CC   and HLA compatibility, and in genetic fingerprinting. Compared with
CC   known methods using MS, this process requires fewer specific reagents
CC   and is better suited to automation. Extended primers are shorter;
CC   primer annealing is more efficient and the process allows detection of
CC   many sequences simultaneously. The present sequence represent an
CC   oligonucleotide used in an example from the present invention.
SQ   Sequence   88 BP;    56 A;    5 C;    7 G;    20 T;


  Query Match             0.2%;  Score 63.4;  DB 1;  Length 88;
  Best Local Similarity   98.5%;  Pred. No. 0.2;
  Matches   64;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy  18514 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTTGAGACAGAG 18573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTTGAGACAGAG 6

Qy  18574 TCTGA 18578
          ||| |
Db      5 TCTCA 1


RESULT   6
T24892/c
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;


  Query Match             0.2%;  Score 62.4;  DB 1;  Length 100;
  Best Local Similarity   75.8%;  Pred. No. 0.28;
  Matches   75;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy  22621 TTTATTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGGTAGAATGCGGTGGTGTGATC 22680
          ||| ||| ||| | ||||||| ||||||||||||||||||   || |||   ||||  |||
Db    100 TTTGTTTGTTTCAAACAGAGTGTCACTCTGTCACCCAGGCNGGAGTGCAANGGTGCAATC 41

Qy  22681 TTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAAT 22719
          |  ||| | |||||  || ||||||  ||||||||| ||
Db     40 TCAGCTNATTGCAAATTCTGCCTCCCAGGTTCAAGCGAT 2


RESULT   7
T26213
ID   T26213 standard; cDNA to mRNA; 103 BP.
AC   T26213;
DT   13-NOV-1996  (first entry)
DE   Human gene signature HUMGS08452.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
```

```
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2029; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    33 A;    21 C;    25 G;    23 T;


  Query Match             0.2%;  Score 62.6;  DB 1;  Length 103;
  Best Local Similarity   76.2%;  Pred. No. 0.27;
  Matches   77;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy   1313 ATGACTTGAGCTCCGGAGGGGGAGGTTGCAGTGAGCCCAGATTACACCACTGCACTCCAG 1372
           || |||||||  | ||||   |  ||| ||||||||   |||  ||||||||||||||||
Db      2 ATCACTTGAGTCCAGGAGTTTGGTGTTACAGTGAGCTATGATGGCACCACTGCACTCCAG 61

Qy   1373 CCTGGGTGAAAGAGCGAGAATCTGTCTTAAAAAAAAAAAAA 1413
           ||||||  | ||||  ||||  |||||| || |||||||||
Db     62 CCTGGGCCACAGAGTAAGAACATGTCTTTAAGAAAAAAAAA 102


RESULT   8
T25009
ID   T25009 standard; cDNA to mRNA; 108 BP.
AC   T25009;
DT   07-NOV-1996  (first entry)
DE   Human gene signature HUMGS07131.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1748; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    34 A;    31 C;    26 G;    15 T;


  Query Match             0.2%;  Score 62.4;  DB 1;  Length 108;
  Best Local Similarity   73.6%;  Pred. No. 0.29;
  Matches   78;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy  27600 ATCACTTGAACTCAGGAGGCAGAGGTTGTAGTGAGCTGAGATCGCACCACTGCACTCCAG 27659
          ||| | ||| | || |||||  ||| || |||||||   | || | |||||| | |||||
Db      2 ATCGCCTGAGCCCATGAGGCCAAGGCTGCAGTGAGCCATGGTCACGCCACTGNATTCCAG 61

Qy  27660 CCTGGGTGACAGAGCAAGACTCCATTTAAAAAAAAAATAATAATAA 27705
          |||| ||||||||||||||| |  || |||| || || || |  ||
Db     62 CCTGAGTGACAGAGCAAGACCCTGTTGAAAACAACAACAACANCAA 107


RESULT   9
T21566/c
ID   T21566 standard; cDNA to mRNA; 87 BP.
AC   T21566;
DT   03-AUG-1996  (first entry)
DE   Human gene signature HUMGS02944.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 914; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   87 BP;    35 A;    21 C;    16 G;    13 T;


  Query Match             0.2%;  Score 61.6;  DB 1;  Length 87;
  Best Local Similarity   81.4%;  Pred. No. 0.37;
  Matches   70;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  15085 TGCCTGGCTAATTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGG 15144
          | | ||||| |||| |||||| || ||| ||||||||||||||||||||   | ||||||
```

```
Db     86 TTCTTGGCTNATTTCTGTATTTTTTGTAAAGATGGGGTTTCGCCATGTTCCTCGGGCTGG 27

Qy  15145 TCTCGAACTCCTGGCCTCAAGCGATC 15170
          | |  |||||||||  ||||||||||
Db     26 TTTTAAACTCCTGGGNTCAAGCGATC 1


RESULT  10
T20927
ID   T20927 standard; cDNA to mRNA; 103 BP.
AC   T20927;
DT   24-JUL-1996  (first entry)
DE   Human gene signature HUMGS02180.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 758-759; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    22 A;    27 C;    21 G;    31 T;


  Query Match             0.2%;  Score 61.8;  DB 1;  Length 103;
  Best Local Similarity   75.8%;  Pred. No. 0.35;
  Matches   75;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy  15033 ATTCTCCTGCCTCAGCCTCCCAACTAGCTGGGACCACAGGTGGGTGCCACCATGCCTGGC 15092
          || ||||  | ||  |||||||| |||||| | | ||||||| |||||||||| |  ||
Db      2 ATCCTCCCACTTCCACCTCCCAAGTAGCTGTGGCTACAGGTGTGTGCCACCATGTCCAGC 61

Qy  15093 TAATTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATG 15131
          | ||||| ||||| | ||||| ||  |  |||| |||||
Db     62 TGATTTTNGTATTTTNAGTAGGGACAGTATTTCTCCATG 100


RESULT  11
T26828/c
ID   T26828 standard; cDNA to mRNA; 108 BP.
AC   T26828;
DT   14-NOV-1996  (first entry)
DE   Human gene signature HUMGS09078.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2182; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    18 A;    33 C;    23 G;    28 T;


  Query Match             0.2%;  Score 61.4;  DB 1;  Length 108;
  Best Local Similarity   79.8%;  Pred. No. 0.4;
  Matches   71;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  16705 AATCAATAACTAGGCTGGGCGTGATGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCC 16764
          ||  |||||    || ||||||| ||||||| |||||||| ||||||||| |||||||||
Db     90 AAGAAATAAAACAGCCGGGCGTGGTGGCTCATGCCTGTAANCCCAGCACTATGGGAGGCC 31

Qy  16765 AAGGCGGACAGATCACGAGGTCAGGAGTT 16793
            |  ||| | ||| |||||||||||||| |
Db     30 GANACGGGCGGATGACGAGGTCAGGAGAT 2


RESULT  12
T25854
ID   T25854 standard; cDNA to mRNA; 91 BP.
AC   T25854;
DT   22-OCT-1996  (first entry)
DE   Human gene signature HUMGS08084.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1944; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
```

```
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   91 BP;    18 A;    22 C;    28 G;    18 T;


  Query Match             0.2%;  Score 61;  DB 1;  Length 91;
  Best Local Similarity   78.7%;  Pred. No. 0.46;
  Matches   70;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy  16911 ATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGAGATCGCGCCACTGCATTCCAG 16970
          ||||||||| || |||||||||| | ||  |||||||||||||| || |  ||||  |||||
Db      2 ATCACTTGAGCCTAGGAGGCAGNGGTTCAAGTGAGCTGAGATGGCACTCCTGCGCTCCAG 61

Qy  16971 CCTGGGAGACAGAGCGAGACTCCGTCTCA 16999
          ||| || ||||| | ||||  | ||||||
Db     62 CCTNGGTGACAGCGTGAGANNCTGTCTCA 90


RESULT  13
T26828
ID   T26828 standard; cDNA to mRNA; 108 BP.
AC   T26828;
DT   14-NOV-1996  (first entry)
DE   Human gene signature HUMGS09078.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2182; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    18 A;    33 C;    23 G;    28 T;


  Query Match             0.2%;  Score 61;  DB 1;  Length 108;
  Best Local Similarity   73.8%;  Pred. No. 0.46;
  Matches   76;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy    509 GATCTCTTGACCTTGTGATCCACCCGCCTCAGCCTCCCAAAGTGCCAGGATTACAGGCAT 568
          |||||| |||||| || |||| |||| || |||||||| |||||  || ||||||||||
Db      1 GATCTCCTGACCTCGTCATCCGCCCGTNTCGGCCTCCCATAGTGCTGGGNTTACAGGCAT 60

Qy    569 GAGCCACCGTGCCCAGCCTCTTTTTCTTTTCTTATAAGACAAG 611
          ||||||||  |||| ||   ||| | | || | |    ||| |
Db     61 GAGCCACCACGCCCGGCTGTTTTATTTCTTATAACTGTACAGG 103


RESULT  14
T22572/c
ID   T22572 standard; cDNA to mRNA; 93 BP.
AC   T22572;
DT   01-OCT-1996  (first entry)
DE   Human gene signature HUMGS04188.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1159; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   93 BP;    22 A;    23 C;    24 G;    22 T;


  Query Match             0.2%;  Score 60.6;  DB 1;  Length 93;
  Best Local Similarity   77.4%;  Pred. No. 0.52;
  Matches   72;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy  19059 TTTTTACGCAGAGTCCTACTCTGTCACCCAGGCTGGAGTGCAGCGGTGTGAACATGGCTC 19118
          ||| ||  ||| ||| | |    | ||| | |||||||||||| ||||||| ||| ||||
Db     93 TTTGTAAACAGGGTCTTGCATNATAACCAANGCTGGAGTGCAGTGGTGTGACCATAGCTC 34

Qy  19119 ACTGCAGCCTCAACCTACTGAGCTCAAGCAATC 19151
          ||||||||||||||| ||  |||||||  |||
Db     33 ACTGCAGCCTCAACCTCCTCGGCTCAAGGGATC 1


RESULT  15
T26213/c
ID   T26213 standard; cDNA to mRNA; 103 BP.
AC   T26213;
DT   13-NOV-1996  (first entry)
DE   Human gene signature HUMGS08452.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
```

PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.2%; Score 60.4; DB 1; Length 103;
Best Local Similarity 74.5%; Pred. No. 0.56;
Matches 76; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy  19050 TTTTTTTTTTTTTTACGCAGAGTCCTACTCTGTCACCCAGGCTGGAGTGCAGCGGTGTGA 19109
          ||||||||| ||  |  ||   || ||||||||  |||||||||||||||| ||||  |
Db    102 TTTTTTTTTCTTAAAGACATGTTCTTACTCTGTGGCCCAGGCTGGAGTGCAGTGGTGCCA 43

Qy  19110 ACATGGCTCACTGCAGCCTCAACCTACTGAGCTCAAGCAATC 19151
            ||| |||||||| | |  ||| || |||  ||||||  |||
Db     42 TCATAGCTCACTGTAACACCAAACTCCTGGACTCAAGTGATC 1
```

Search completed: June 17, 2000, 11:43:03
Job time: 254694 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 15:09:23 ; Search time 13753.1 Seconds
(without alignments)
8546.953 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTTAGAGGAAAAAAAGCA.....TAGATAAAACGTTGTTCCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query

	No.	Score	Match	Length	DB	ID	Description
	1	89	0.3	105	105	AQ282107	AQ282107 RPCI11-94
c	2	87.2	0.3	108	84	B65160	B65160 CIT-HSP-201
	3	87.2	0.3	109	84	B17434	B17434 345K2.TVB C
	4	87.2	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
c	5	86.8	0.3	106	37	AA703692	AA703692 ag81a10.r
c	6	86.8	0.3	109	24	N25299	N25299 yw52c09.s1
c	7	86.4	0.3	107	33	AA385808	AA385808 EST99495
c	8	85	0.3	101	35	AA583697	AA583697 nn58f10.s
c	9	85.2	0.3	106	63	AI991750	AI991750 wt48e01.x
c	10	85	0.3	109	22	H11143	H11143 ym09c06.r1
	11	85	0.3	110	30	AA244245	AA244245 nc07a04.s
	12	84.4	0.3	106	30	AA250812	AA250812 zs06a05.s
c	13	84.6	0.3	107	35	AA565533	AA565533 nk42b11.s
	14	83.8	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
	15	84	0.3	109	103	AQ200347	AQ200347 RPCI11-43
c	16	83.2	0.3	106	94	AQ046231	AQ046231 RPCI11-36
c	17	83.4	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
c	18	83.4	0.3	110	30	AA244245	AA244245 nc07a04.s
c	19	82.8	0.3	102	36	AA654562	AA654562 nt75f10.s
	20	82.8	0.3	107	24	H67040	H67040 yu68c01.r1
	21	82.8	0.3	110	106	AQ386882	AQ386882 RPCI11-13
c	22	82.2	0.3	103	108	AQ535244	AQ535244 RPCI-11-3
c	23	82.2	0.3	106	94	AQ062963	AQ062963 CIT-HSP-2
c	24	82.4	0.3	109	30	AA243009	AA243009 zr25h02.s
c	25	81.8	0.3	101	39	AA835205	AA835205 ak64h01.s
	26	82	0.3	106	105	AQ282340	AQ282340 RPCI11-80
	27	81.6	0.3	104	29	AA129957	AA129957 zn86h04.r
c	28	81	0.3	106	44	AI249096	AI249096 qh73g09.x
c	29	81.2	0.3	110	106	AQ386882	AQ386882 RPCI11-13
	30	80.6	0.3	103	108	AQ584425	AQ584425 RPCI-11-4
c	31	80.6	0.3	104	105	AQ321855	AQ321855 RPCI11-11
c	32	80.4	0.3	106	38	AA812141	AA812141 ob48h02.s
c	33	80.4	0.3	106	106	AQ414071	AQ414071 RPCI-11-1
c	34	80	0.3	105	30	AA218889	AA218889 zq15d04.s
	35	80	0.3	107	103	AQ240182	AQ240182 CIT-HSP-2
c	36	79.6	0.3	102	84	B48088	B48088 RPCI11-4N6.
	37	79.8	0.3	107	62	AI933497	AI933497 wm74d02.x
	38	79.8	0.3	108	84	B65160	B65160 CIT-HSP-201
c	39	79.8	0.3	110	39	AA897366	AA897366 am06h02.s
c	40	79.2	0.3	100	30	AA252633	AA252633 zq43g05.r
	41	79.4	0.3	105	28	AA078003	AA078003 7H12D08 C
c	42	79.4	0.3	105	105	AQ276193	AQ276193 CITBI-E1-
c	43	79.4	0.3	107	24	H67040	H67040 yu68c01.r1
	44	79.6	0.3	110	109	AQ634950	AQ634950 RPCI-11-4
c	45	79	0.3	103	35	AA570438	AA570438 nk63g02.s

ALIGNMENTS

```
RESULT   1
AQ282107
LOCUS       AQ282107      105 bp    DNA             GSS       27-APR-1999
DEFINITION  RPCI11-94B21.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-94B21,
            genomic survey sequence.
ACCESSION   AQ282107
VERSION     AQ282107.1  GI:3907976
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 105)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .105
                     /organism="Homo sapiens"
                     /db_xref="GDB:7535756"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-94B21"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       26 a     31 c     30 g     18 t
ORIGIN


  Query Match             0.3%;  Score 89;  DB 105;  Length 105;
  Best Local Similarity   90.5%;  Pred. No. 0.21;
  Matches   95;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16731 GCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGA 16790
          ||||||||||||||||||||||| |||||||||||| ||   |||||||||| || ||
Db      1 GCTCACGCCTGTAATCCCAGCACTCTGGGAGGCCAAGGTGGGTGGATCACGAGGGCATGA 60

Qy  16791 GTTTGAGACCAGCCTGACCAACATGGTGAAACCCTGTCTCTACTA 16835
          ||  |||||||||||||||||||||||||||||| ||||||||||
Db     61 GTACGAGACCAGCCTGACCAACATGGTGAAACCCCGTCTCTACTA 105


RESULT   2
B65160/c
LOCUS       B65160        108 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
            genomic survey sequence.
ACCESSION   B65160
VERSION     B65160.1  GI:2639138
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2017G2.TFB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
```

```
                     /organism="Homo sapiens"
                     /db_xref="GDB:7043860"
                     /db_xref="taxon:9606"
                     /clone="2017G2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       26 a     27 c     34 g     21 t
ORIGIN


  Query Match             0.3%;  Score 87.2;  DB 84;  Length 108;
  Best Local Similarity   88.0%;  Pred. No. 0.33;
  Matches   95;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  22662 AGAATGCGGTGGTGTGATCTTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCAATTC 22721
          ||  ||| ||||| |||||||||||||||||||||||||| ||||| |||||||||  ||||
Db    108 AGTGTGCAGTGGTATGATCTTGGCTCACTGCAACCTCCACCTCCCGGGTTCAAGAGATTC 49

Qy  22722 CCATGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACATACCACCA 22769
          | |||||||||||| |||||||||||||||||||| ||| ||||||
Db     48 TCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCATGCCACCA 1


RESULT   3
B17434
LOCUS       B17434        109 bp    DNA             GSS       04-JUN-1998
DEFINITION  345K2.TVB CIT978SKA1 Homo sapiens genomic clone A-345K02, genomic
            survey sequence.
ACCESSION   B17434
VERSION     B17434.1  GI:2125183
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: 345K02.TP 345K02.TPB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="A-345K02"
                     /clone_lib="CIT978SKA1"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A1"
BASE COUNT       24 a     30 c     31 g     24 t
ORIGIN


  Query Match             0.3%;  Score 87.2;  DB 84;  Length 109;
  Best Local Similarity   88.0%;  Pred. No. 0.33;
  Matches   95;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  27417 GGCTCATACCTGTAATCCCAGCACTTTGGGAGACTGAGGTGGGTGGATCACTTGAGGTCA 27476
          |||||||||| |||||| |||||||||||| |||| ||||| ||||||| |||||||
Db      2 GGCTCATACCTATAATCCTAGCACTTTGGGAGGCTGATGTGGGCGGATCACCTGAGGTCG 61

Qy  27477 GGAGTTCAAAACCAGCCTGGCCAACATGGTGAAAACCCATCTCTACTA 27524
          ||||||| | |||||||||||| |||||||||| ||| |||| ||||
Db     62 GGAGTTCGAGACCAGCCTGGCCACCATGGTGAAACCCCGTCTCAACTA 109


RESULT   4
AQ028426
LOCUS       AQ028426      109 bp    DNA             GSS       30-JUN-1998
DEFINITION  CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
            genomic survey sequence.
ACCESSION   AQ028426
VERSION     AQ028426.1  GI:3268648
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2313G15"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       19 a     36 c     25 g     29 t
ORIGIN


  Query Match             0.3%;  Score 87.2;  DB 94;  Length 109;
  Best Local Similarity   88.0%;  Pred. No. 0.33;
  Matches   95;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  22365 TTTTTTCTGAAATGGAGTCTCACTCTGTTGCCCAGGCTGGAGTACAGTGGCACAATCTTG 22424
          | |||||||| | ||| |||||||||||  |||||||||||| |||||||||| |||
Db      2 TGTTTTCTGAGACGGACTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCACAGTCTGA 61

Qy  22425 GTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGAGTCTCCTGACTCA 22472
          | |||||||||||||||||||||||||||||||| ||||||| ||||
Db     62 GCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCA 109


RESULT   5
AA703692/c
LOCUS       AA703692      106 bp    mRNA            EST       24-DEC-1997
DEFINITION  ag81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:1140858 5' similar to contains Alu repetitive element;, mRNA
            sequence.
```

```
ACCESSION   AA703692
VERSION     AA703692.1  GI:2713610
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1397630.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 53.
FEATURES             Location/Qualifiers
     source          1. .106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1140858"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     29 c     29 g     29 t
ORIGIN
```

```
  Query Match             0.3%;  Score 86.8;  DB 37;  Length 106;
  Best Local Similarity   88.7%;  Pred. No. 0.37;
  Matches   94;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  16734 CACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTT 16793
          ||||||||||||||||||||||||||||||  ||||||  |||||||||||||||||||| |
Db    106 CACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCAGATCACGAGGTCAGGAGAT 47

Qy  16794 TGAGACCAGCCTGACCAACATGGTGAAACCCTGTCTCTACTAACAA 16839
          ||||||| |||| | |||| |||||||||  ||||||||||| ||
Db     46 CGAGACCATCCTGGCTAACACGGTGAAACCTCGTCTCTACTAAAAA 1
```

```
RESULT   6
N25299/c
LOCUS       N25299        109 bp    mRNA            EST       28-DEC-1995
DEFINITION  yw52c09.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
            IMAGE:255856 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   N25299
VERSION     N25299.1  GI:1139449
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     On Apr 14, 1993 this sequence version replaced gi:837394.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 307.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="GDB:3866265"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:255856"
                     /clone_lib="Weizmann Olfactory Epithelium"
                     /sex="Female"
                     /tissue_type="olfactory epithelium"
                     /dev_stage="35 year old"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: nose; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Olfactory epithelium, normal. Average insert
                     size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N.
                     Walker, D. Lancet, Weizmann Institute of Science. ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       13 a     34 c     24 g     35 t      3 others
ORIGIN
```

```
  Query Match             0.3%;  Score 86.8;  DB 24;  Length 109;
  Best Local Similarity   86.2%;  Pred. No. 0.37;
  Matches   94;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  16901 AGGCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGAGATCGCGCCAC 16960
          ||||||||||||  |  |||||  |||||||||| |||||||||| |||||| ||||||
Db    109 AGGCAGGAGAATGGCAGGAACCTGGGAGGCAGAGTTTGCAGTGAGCAGAGATCACGCCAC 50

Qy  16961 TGCATTCCAGCCTGGGAGACAGAGCGAGACTCCGTCTCAAAATTAAAAA 17009
          |||| ||||||||||||||||||||||   ||||||||||||  |||||
Db     49 TGCACTCCAGCCTGGGAGACAGAGCGAAGNTCCGTCTCAAAANNAAAAA 1
```

```
RESULT   7
AA385808/c
LOCUS       AA385808      107 bp    mRNA            EST       21-APR-1997
DEFINITION  EST99495 Thyroid Homo sapiens cDNA 5' end similar to EST containing
            Alu repeat, mRNA sequence.
ACCESSION   AA385808
VERSION     AA385808.1  GI:2038127
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
```

```
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., ,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   12140200
COMMENT     On Jan 25, 1995 this sequence version replaced gi:637865.
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):189984"
                     /db_xref="taxon:9606"
                     /clone_lib="Thyroid"
                     /dev_stage="adult"
                     /note="Organ: thyroid gland; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoI"
BASE COUNT       16 a     34 c     28 g     26 t      3 others
ORIGIN


  Query Match             0.3%;  Score 86.4;  DB 33;  Length 107;
  Best Local Similarity   86.9%;  Pred. No. 0.41;
  Matches   93;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16892 GGGACACTGAGGCAGGAGAATCACTTGAACCCAGGAGGCAGAGATTGCAGTGAGCTGAGA 16951
          ||||  |||||||||||||||   | |||||||||| |||||||| ||||||||||||||||
Db    107 GGGAGNCTGAGGCAGGAGAACGNCGTGAACCCAGAAGGCAGAGCTTGCAGTGAGCTGAGA 48

Qy  16952 TCGCGCCACTGCATTCCAGCCTGGGAGACAGAGCGAGACTCCGTCTC 16998
          || |||||||||| |||| ||||||  |||||||||||||| |||||
Db     47 TCTCGCCACTGCACTCCAACCTGGGTNACAGAGCGAGACTCTGTCTC 1


RESULT   8
AA583697/c
LOCUS       AA583697      101 bp    mRNA            EST       26-SEP-1997
DEFINITION  nn58f10.s1 NCI_CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1088107 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA583697
VERSION     AA583697.1  GI:2368306
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 101)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Apr 14, 1993 this sequence version replaced gi:692704.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1890   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 93.
FEATURES             Location/Qualifiers
     source          1. .101
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1088107"
                     /clone_lib="NCI_CGAP_Kid6"
                     /sex="mixed"
                     /tissue_type="kidney tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: kidney; Vector: Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3' Average insert size:  1.0 kb."
BASE COUNT       21 a     29 c     28 g     23 t
ORIGIN


  Query Match             0.3%;  Score 85;  DB 35;  Length 101;
  Best Local Similarity   90.1%;  Pred. No. 0.61;
  Matches   91;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  16733 TCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGT 16792
          ||||||||||||||||||||||||||||||||||||||||| | |||||| ||||| |||||
Db    101 TCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGCGGATCACAAGGTCTGGAGT 42

Qy  16793 TTGAGACCAGCCTGACCAACATGGTGAAACCCTGTCTCTAC 16833
          | |||| ||||||| |||||||||||||| || || |||||
Db     41 TCGAGATCAGCCTGGCCAACATGGTGAAATCCCGTTTCTAC 1


RESULT   9
AI991750/c
LOCUS       AI991750      106 bp    mRNA            EST       08-SEP-1999
DEFINITION  wt48e01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2510712 3'
            similar to contains Alu repetitive element;contains element LTR8
            repetitive element ;, mRNA sequence.
ACCESSION   AI991750
VERSION     AI991750.1  GI:5838578
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Dec 20, 1995 this sequence version replaced gi:1133359.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -40UP from Gibco
```

```
          High quality sequence stop: 62.
FEATURES             Location/Qualifiers
     source          1. .106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2510712"
                     /clone_lib="NCI_CGAP_Pan1"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
BASE COUNT       24 a     23 c     22 g     37 t
ORIGIN


  Query Match             0.3%;  Score 85.2;  DB 63;  Length 106;
  Best Local Similarity   87.7%;  Pred. No. 0.56;
  Matches   93;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   1144 AATCCTAGCACTTTGGGAAGCCGAGGCGGGCAGATCACCTGAGGTCAGGAGTTTGAGACC 1203
           ||||| |||||||||||| || ||||| || ||||||||||||||||||||||||||||
Db    106 AATCCCAGCACTTTGGGAGGCTGAGGCTGGTAGATCACCTGAGGTCAGGAGTTTGAGACT 47

Qy   1204 AGCCTGACCAACATGGCAAAACCTCATCTCTACTAAAAATACAAAA 1249
           | |||  |||||||| |||||||||||||| |||||| | ||||
Db     46 AACCTTGCCAACATGGTAAAACCTCATCTCTATTAAAAAAAAAAA 1


RESULT  10
H11143/c
LOCUS       H11143        109 bp    mRNA            EST       26-JUN-1995
DEFINITION  ym09c06.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:47310 5' similar to contains Alu repetitive element;contains
            MER22 repetitive element ;, mRNA sequence.
ACCESSION   H11143
VERSION     H11143.1  GI:875963
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     On May 5, 1995 this sequence version replaced gi:798506.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1316
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Putative full length read
            Insert Length: 1316   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 363.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="GDB:419851"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:47310"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       27 a     28 c     36 g     18 t
ORIGIN


  Query Match             0.3%;  Score 85;  DB 22;  Length 109;
  Best Local Similarity   86.2%;  Pred. No. 0.59;
  Matches   94;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    310 TGAGACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGACGCTATCTCGGCTCACT 369
          ||||| || |||||||||||||||||||||||||||||||| | | ||||| |||||||
Db    109 TGAGAAGGCGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCTCGATCTCAGCTCACT 50

Qy    370 GCACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGA 418
          | | |  |  |||||||||||||| |||||||||||| ||||||| ||
Db     49 GGAACCTCCGCTCCCGGGTTCAAGTGATTCTCCTGCCTTAGCCTCCTGA 1


RESULT  11
AA244245
LOCUS       AA244245      110 bp    mRNA            EST       20-AUG-1997
DEFINITION  nc07a04.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1007406
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA244245
VERSION     AA244245.1  GI:1875104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 24, 1995 this sequence version replaced gi:634306.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 90.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1007406"
                     /clone_lib="NCI_CGAP_Pr1"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from
```

5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN

Query Match 0.3%; Score 85; DB 30; Length 110;
Best Local Similarity 85.5%; Pred. No. 0.59;
Matches 94; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 22363 TGTTTTTTCTGAAATGGAGTCTCACTCTGTTGCCCAGGCTGGAGTACAGTGGCACAATCT 22422
Db 1 TTTTTTTTTTGAGATGGAGTCTTGATCTGTTGCCCAGGCTGGAGTGCAGTGGCAGANTCT 60

Qy 22423 TGGTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGAGTCTCCTGACTCA 22472
Db 61 TGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTTCTGCCTCA 110

RESULT 12
AA250812
LOCUS AA250812 106 bp mRNA EST 15-AUG-1997
DEFINITION zs06a05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:684368 3' similar to contains Alu repetitive element;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AA250812
VERSION AA250812.1 GI:1885774
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407356.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 537 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 87.
FEATURES Location/Qualifiers
source 1. .106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:684368"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 20 a 28 c 31 g 27 t
ORIGIN

Query Match 0.3%; Score 84.4; DB 30; Length 106;
Best Local Similarity 89.2%; Pred. No. 0.7;
Matches 91; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 15110 GTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCCTGGCCTCAAGCGAT 15169
Db 1 GTAGAGACGGGGTTTCACCATGTTGGCCAGGCCGGTCTCAAACTCTTGACCTCAGGTGAT 60

Qy 15170 CCACTTGCCTTGGCCTCCCAAAGTGCTAGGATTACAGGCATT 15211
Db 61 CCACTTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTT 102

RESULT 13
AA565533/c
LOCUS AA565533 107 bp mRNA EST 08-SEP-1997
DEFINITION nk42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AA565533
VERSION AA565533.1 GI:2337172
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1661 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
FEATURES Location/Qualifiers
source 1. .107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI_CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Bulk germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCACGAG 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 22 a 34 c 26 g 25 t
ORIGIN

Query Match 0.3%; Score 84.6; DB 35; Length 107;
Best Local Similarity 86.9%; Pred. No. 0.66;
Matches 93; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 16864 TGGCATCTGCCTGTAGTCCCAGCTACTCGGGACACTGAGGCAGGAGAATCACTTGAACCC 16923

```
          ||| | |||||||| |||||||||||| ||| ||||||||| ||||||||||||||||
Db    107 TGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCACGAGAATCACTTGAACCT 48

Qy  16924 AGGAGGCAGAGATTGCAGTGAGCTGAGATCGCGCCACTGCATTCCAG 16970
          |||||||||| |||||||||||||||| | |||||||| |||||
Db     47 GGGAGGCAGAGCTTGCAGTGAGCTGAGATTGAGCCACTGCACTCCAG 1


RESULT  14
AQ535244
LOCUS       AQ535244      103 bp    DNA             GSS       18-MAY-1999
DEFINITION  RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-317H22, genomic survey sequence.
ACCESSION   AQ535244
VERSION     AQ535244.1  GI:4846934
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7621533"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-317H22"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       31 a     27 c     27 g     18 t
ORIGIN


  Query Match             0.3%;  Score 83.8;  DB 108;  Length 103;
  Best Local Similarity   88.3%;  Pred. No. 0.83;
  Matches   91;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  17732 CCAGCACTTTTGGAGGCCGAGGCGTGCAGATCACTTGAGGACAGGAGTTCAAGACCAGCC 17791
          |||||||||| ||||||| || || ||||||||||||||| ||||||||| |||||||||
Db      1 CCAGCACTTTGGGAGGCCAAGACGGGCAGATCACTTGAGGTCAGGAGTTCGAGACCAGCC 60

Qy  17792 TGGCTAACATGGCAAAACCCCATCTCTACTAAAAATACAAAAA 17834
          |||| |||||||  ||||||| ||||| ||| ||||||||||
Db     61 TGGCCAACATGGTGAAACCCCGTCTCTGCTATAAATACAAAAA 103


RESULT  15
AQ200347
LOCUS       AQ200347      109 bp    DNA             GSS       20-APR-1999
DEFINITION  RPCI11-43B21.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-43B21,
            genomic survey sequence.
ACCESSION   AQ200347
VERSION     AQ200347.1  GI:3612546
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
  TITLE     Use of human BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: RPCI11-43B21.TK
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="GDB:7516172"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-43B21"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       27 a     31 c     29 g     22 t
ORIGIN


  Query Match             0.3%;  Score 84;  DB 103;  Length 109;
  Best Local Similarity   86.1%;  Pred. No. 0.77;
  Matches   93;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  17783 AGACCAGCCTGGCTAACATGGCAAAACCCCATCTCTACTAAAAATACAAAAATTAACCAG 17842
          | |||||||||||| |||| ||  |||||||||||||||| ||| ||||  ||||| ||||
Db      2 ATACCAGCCTGGCGAACACGGTGAAACCCCATCTCTACTGAAAGTACATTAATTAGCCAG 61

Qy  17843 GCGTGGTGGTGCACGCCTGTAATCCCAGCTACTCTGGAGGCTGAGGCA 17890
          | ||||||| |||||||||| |||||||||||||| ||||||||| |||
Db     62 GTGTGGTGGCGCACGCCTGTGATCCCAGCTACTCGGGAGGCTGACGCA 109


Search completed: June 17, 2000, 02:34:25
Job time: 223032 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 20:15:34 ; Search time 593.69 Seconds
(without alignments)
6349.604 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTTAGAGGAAAAAAAGCA.....TAGATAAAACGTTGTTCCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	82.6	0.3	105	4	US-08-481-658B-65	Sequence 65, Appl
c	2	82.6	0.3	105	4	US-08-477-504A-65	Sequence 65, Appl
c	3	82.6	0.3	105	4	US-08-486-756A-65	Sequence 65, Appl
c	4	82.6	0.3	105	4	US-08-485-862B-65	Sequence 65, Appl
c	5	82.6	0.3	105	5	US-08-787-739-65	Sequence 65, Appl
c	6	68.6	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
c	7	68.6	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
c	8	68.6	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
c	9	68.6	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
	10	64.8	0.2	105	4	US-08-481-658B-65	Sequence 65, Appl
	11	64.8	0.2	105	4	US-08-477-504A-65	Sequence 65, Appl
	12	64.8	0.2	105	4	US-08-486-756A-65	Sequence 65, Appl
	13	64.8	0.2	105	4	US-08-485-862B-65	Sequence 65, Appl
	14	64.8	0.2	105	5	US-08-787-739-65	Sequence 65, Appl
	15	59.4	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
	16	59.4	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
	17	59.4	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
	18	59.4	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
c	19	57.8	0.2	78	3	US-08-454-557C-70	Sequence 70, Appl
c	20	57.8	0.2	78	4	US-08-340-426D-70	Sequence 70, Appl
c	21	57.8	0.2	78	4	US-08-450-673C-70	Sequence 70, Appl
c	22	57.8	0.2	78	6	PCT-US95-17111A-70	Sequence 70, Appl
	23	56.6	0.2	92	1	US-08-222-177A-430	Sequence 430, App
c	24	56	0.2	85	3	US-08-454-557C-92	Sequence 92, Appl
c	25	56	0.2	85	4	US-08-340-426D-92	Sequence 92, Appl
c	26	56	0.2	85	4	US-08-450-673C-92	Sequence 92, Appl
c	27	56	0.2	85	6	PCT-US95-17111A-92	Sequence 92, Appl
	28	54	0.2	60	1	US-08-222-177A-244	Sequence 244, App
	29	53.6	0.2	91	1	US-08-222-177A-166	Sequence 166, App
	30	52.4	0.2	84	3	US-08-454-557C-91	Sequence 91, Appl
	31	52.4	0.2	84	4	US-08-340-426D-91	Sequence 91, Appl
	32	52.4	0.2	84	4	US-08-450-673C-91	Sequence 91, Appl
	33	52.4	0.2	84	6	PCT-US95-17111A-91	Sequence 91, Appl
c	34	50.4	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
c	35	50.4	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
c	36	50.4	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
c	37	50.4	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
	38	50	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
	39	50	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
	40	50	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
	41	50	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
	42	50	0.2	83	4	US-08-481-658B-66	Sequence 66, Appl
	43	50	0.2	83	4	US-08-477-504A-66	Sequence 66, Appl
	44	50	0.2	83	4	US-08-486-756A-66	Sequence 66, Appl
	45	50	0.2	83	4	US-08-485-862B-66	Sequence 66, Appl

ALIGNMENTS

```
RESULT   1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match              0.3%;  Score 82.6;  DB 4;  Length 105;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-09;
```

```
  Matches   91;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16744 ATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGAGACCAGC 16803
          |||||||||||||||||||| |||| |   |||||| |||||||||||||||| ||||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  16804 CTGACCAACATGGTGAAACCCTGTCTCTACTAACAAAATACAAAA 16848
          ||| |||| |||||||||||||||||||||||  |  || ||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAAA 1


RESULT   2
US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.3%;  Score 82.6;  DB 4;  Length 105;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-09;
  Matches   91;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16744 ATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGAGACCAGC 16803
          |||||||||||||||||||| |||| |   |||||| |||||||||||||||| ||||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  16804 CTGACCAACATGGTGAAACCCTGTCTCTACTAACAAAATACAAAA 16848
          ||| |||| |||||||||||||||||||||||  |  || ||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAAA 1


RESULT   3
US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.3%;  Score 82.6;  DB 4;  Length 105;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-09;
  Matches   91;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16744 ATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGAGACCAGC 16803
          |||||||||||||||||||| |||| |   |||||| |||||||||||||||| ||||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  16804 CTGACCAACATGGTGAAACCCTGTCTCTACTAACAAAATACAAAA 16848
          ||| |||| |||||||||||||||||||||||  |  || ||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAAA 1


RESULT   4
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
```

```
;       CITY:  Tiburon
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94920
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,862B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;       APPLICATION NUMBER:  US 08/260,190
;       FILING DATE:  15-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.3D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.3%;  Score 82.6;  DB 4;  Length 105;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-09;
  Matches   91;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16744 ATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGAGACCAGC 16803
          |||||||||||||||||||||| |||| |   |||||| |||||||||||||||||| ||||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  16804 CTGACCAACATGGTGAAACCCTGTCTCTACTAACAAAATACAAAA 16848
          ||| |||| ||||||||||||||||||||||||  |  || ||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAAA 1


RESULT   5
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;     APPLICANT:  Zavada, Jan
;     APPLICANT:  Pastorekova, Silvia
;     APPLICANT:  Pastorek, Jaromir
;     TITLE OF INVENTION:  MN Gene and Protein
;     NUMBER OF SEQUENCES:  96
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Leona L. Lauder
;       STREET:  369 Pine Street, Suite 610
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/787,739
;       FILING DATE:  24-JAN-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,049
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/486,756
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/481,658
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,862
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,863
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/487,077
;       FILING DATE:  07-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-981-2034
;       TELEFAX:  415-981-0332
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-787-739-65


  Query Match             0.3%;  Score 82.6;  DB 5;  Length 105;
  Best Local Similarity   86.7%;  Pred. No. 1.7e-09;
  Matches   91;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  16744 ATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTTTGAGACCAGC 16803
          |||||||||||||||||||||| |||| |   |||||| |||||||||||||||||| ||||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAAGGTCAGGAGTTTGAGAGCAGC 46

Qy  16804 CTGACCAACATGGTGAAACCCTGTCTCTACTAACAAAATACAAAA 16848
          ||| |||| ||||||||||||||||||||||||  |  || ||||
Db     45 CTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAAA 1


RESULT   6
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;     APPLICANT:  de la Monte, Suzanne
;     APPLICANT:  Wands, Jack R.
;     TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;     TITLE OF INVENTION:  of Alzheimer's Disease
;     NUMBER OF SEQUENCES:  121
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;       STREET:  1100 New York Avenue, Suite 600
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-91


  Query Match             0.2%;  Score 68.6;  DB 3;  Length 84;
  Best Local Similarity   89.2%;  Pred. No. 1.8e-06;
  Matches   74;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  16734 CACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTT 16793
          ||||| |||||||||||||||||||||||||  |||||| | ||||||||||||||||||
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  16794 TGAGACCAGCCTGACCAACATGG 16816
            || ||||||||||  |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT   7
US-08-340-426D-91/c
; Sequence 91, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-91


  Query Match             0.2%;  Score 68.6;  DB 4;  Length 84;
  Best Local Similarity   89.2%;  Pred. No. 1.8e-06;
  Matches   74;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  16734 CACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTT 16793
          ||||| |||||||||||||||||||||||||  |||||| | ||||||||||||||||||
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  16794 TGAGACCAGCCTGACCAACATGG 16816
            || ||||||||||  |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT   8
US-08-450-673C-91/c
; Sequence 91, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-91


  Query Match             0.2%;  Score 68.6;  DB 4;  Length 84;
  Best Local Similarity   89.2%;  Pred. No. 1.8e-06;
  Matches   74;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  16734 CACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTT 16793
          ||||| |||||||||||||||||||||||||  |||||| | ||||||||||||||||||
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  16794 TGAGACCAGCCTGACCAACATGG 16816
            || ||||||||||  |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1
```

```
RESULT   9
PCT-US95-17111A-91/c
; Sequence 91, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-91


  Query Match             0.2%;  Score 68.6;  DB 6;  Length 84;
  Best Local Similarity   89.2%;  Pred. No. 1.8e-06;
  Matches   74;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  16734 CACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGACAGATCACGAGGTCAGGAGTT 16793
          ||||| |||||||||||||||||||||||||  |||||| | ||||||||||||||||||
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy  16794 TGAGACCAGCCTGACCAACATGG 16816
            || ||||||||||  |||||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  10
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match             0.2%;  Score 64.8;  DB 4;  Length 105;
  Best Local Similarity   82.1%;  Pred. No. 1.4e-05;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

Qy  15096 TTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCC 15155
          |||||  ||  ||||||||||  ||||||| |||| ||||||||||||| |||| ||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  15156 TGGCCTCAAGCGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGAT 15201
          || |||   | |||||||  |||| ||||||||||||||||| ||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  11
US-08-477-504A-65
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.2%;  Score 64.8;  DB 4;  Length 105;
  Best Local Similarity   82.1%;  Pred. No. 1.4e-05;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

Qy  15096 TTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCC 15155
          |||||  ||  ||||||||||  ||||||| |||| ||||||||||||| |||| ||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  15156 TGGCCTCAAGCGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGAT 15201
          || |||   | |||||||  |||| ||||||||||||||||| ||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  12
US-08-486-756A-65
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.2%;  Score 64.8;  DB 4;  Length 105;
  Best Local Similarity   82.1%;  Pred. No. 1.4e-05;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

Qy  15096 TTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCC 15155
          |||||  ||  ||||||||||  ||||||| |||| ||||||||||||| |||| ||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  15156 TGGCCTCAAGCGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGAT 15201
          || |||   | |||||||  |||| ||||||||||||||||| ||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  13
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,862B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.2%;  Score 64.8;  DB 4;  Length 105;
```

```
  Best Local Similarity   82.1%;  Pred. No. 1.4e-05;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

Qy  15096 TTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCC 15155
          |||||  ||  ||||||||||  |||||| |||| ||||||||||| |||| ||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  15156 TGGCCTCAAGCGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGAT 15201
          || |||   | |||||||  |||| ||||||||||||||||| ||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  14
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-787-739-65


  Query Match             0.2%;  Score 64.8;  DB 5;  Length 105;
  Best Local Similarity   82.1%;  Pred. No. 1.4e-05;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    2;  Gaps    1;

Qy  15096 TTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCGAACTCC 15155
          |||||  ||  ||||||||||  |||||| |||| ||||||||||| |||| ||||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  15156 TGGCCTCAAGCGATCCACTTGCCTTGGCCTCCCAAAGTGCTAGGAT 15201
          || |||   | |||||||  |||| ||||||||||||||||| ||||
Db     62 TGACCT--TGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT  15
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-70


  Query Match             0.2%;  Score 59.4;  DB 3;  Length 78;
  Best Local Similarity   85.7%;  Pred. No. 0.00018;
  Matches   66;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  15081 ACCATGCCTGGCTAATTTTTGTATTATTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGG 15140
          || | |||  |||||| |||||||| ||||||||||||||||||| |||||||   ||||
Db      1 ACAACGCCCAGCTAATATTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTCATCAGG 60

Qy  15141 CTGGTCTCGAACTCCTG 15157
          ||||| ||||||||||
Db     61 CTGGTGTCGAACTCCTG 77
```

Search completed: June 17, 2000, 10:50:11
Job time: 251741 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:40:12 ; Search time 29137.4 Seconds
(without alignments)
-600.987 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTCACTTCTAACCA.....CCATCCAGCTTCTCACTCCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	88.8	0.5	108	10	HSLDLRN2	X05250 Human LDL-r
c	2	87.4	0.5	108	11	HSU67803	U67803 Human small
c	3	84.4	0.5	103	9	HUMALCE221	M87896 Human carci
	4	83	0.5	108	10	HSLDLI12	X05248 Human LDL-r
c	5	82.2	0.5	107	9	HUMALCE162	M87924 Human carci
	6	79.8	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
c	7	79.8	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
	8	79.2	0.4	108	10	HSLDLRN2	X05250 Human LDL-r
c	9	77.8	0.4	108	11	HSU67804	U67804 Human small
c	10	77.2	0.4	91	13	HUMUT8164A	L30244 Human STS U
	11	77.2	0.4	103	13	HS8IC8R	X57789 Human seque
c	12	76.2	0.4	108	11	HSU67808	U67808 Human small
c	13	75	0.4	103	13	HS8IC8R	X57789 Human seque
c	14	74.4	0.4	110	9	HUMALCE43	M87900 Human carci
c	15	74.2	0.4	108	9	HUMD1D03M5	D16965 Human HepG2
	16	74	0.4	103	9	HUMALCE221	M87896 Human carci
c	17	74	0.4	110	11	HSU67807	U67807 Human small
c	18	73.2	0.4	104	9	HUMALCE272	M87899 Human carci
c	19	72	0.4	90	9	HUMLDLRFL	K03555 Human low d
	20	72	0.4	106	13	G32743	G32743 A009P31 Hum
	21	71.2	0.4	108	13	G43535	G43535 WIAF-2393-S
c	22	70.4	0.4	97	9	HUMLDLRA2	M14180 Human low d
c	23	70.2	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
	24	70.2	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
	25	70.2	0.4	110	11	HSU67807	U67807 Human small
	26	69.8	0.4	90	9	HUMLDLRFL	K03555 Human low d
	27	69.8	0.4	104	9	HUMALCE272	M87899 Human carci
	28	69.4	0.4	101	10	S79560	S79560 HRX {intron
	29	69.2	0.4	97	9	HUMLDLRA1	M14178 Human low d
	30	69	0.4	108	11	HSU67803	U67803 Human small
c	31	68.6	0.4	100	13	HUMUT931A	L31299 Human STS U
c	32	68.2	0.4	107	11	HSU67806	U67806 Human small
c	33	68	0.4	108	13	G43535	G43535 WIAF-2393-S
	34	67.8	0.4	108	9	HUMD1D03M5	D16965 Human HepG2
	35	67.2	0.4	80	9	HUMBRKFAE	M36135 Human alpha
	36	67.2	0.4	97	9	HUMLDLRA2	M14180 Human low d
	37	67.2	0.4	99	13	HUMUT7692A	L30306 Human STS U
c	38	67	0.4	97	9	HUMLDLRA1	M14178 Human low d
	39	66.8	0.4	94	9	HUMHGAL	M13479 Human alpha
	40	66.8	0.4	100	11	HSU67848	U67848 Human beta-
c	41	66.6	0.4	101	10	S79560	S79560 HRX {intron
	42	66.6	0.4	108	11	HSU67808	U67808 Human small
c	43	66	0.4	94	9	HUMHGAL	M13479 Human alpha
	44	65.2	0.4	95	13	HUMUT8002B	L30176 Human STS U
	45	65	0.4	98	13	G33095	G33095 EyRP13C9R H

ALIGNMENTS

```
RESULT   1
HSLDLRN2/c
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.5%;  Score 88.8;  DB 10;  Length 108;
  Best Local Similarity   88.9%;  Pred. No. 0.00015;
  Matches   96;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  13470 CTTGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGACTCAGCCTCCCG 13529
          || |||||| |||||||||||||||||||||||||  ||||||||| ||||||||||||
Db    108 CTCGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCG 49

Qy  13530 AGTAGCTGGGATTACAGGCATGCATCACCATGCCTGGGTAATTTTTGT 13577
          |||||||||||||||||||     ||||| |||||| ||||||||||
Db     48 AGTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   2
HSU67803/c
LOCUS       HSU67803      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67803
VERSION     U67803.1  GI:2289917
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu2"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       23 a     39 c     30 g     16 t
ORIGIN


  Query Match             0.5%;  Score 87.4;  DB 11;  Length 108;
  Best Local Similarity   93.8%;  Pred. No. 0.00024;
  Matches   91;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy  12882 GTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 12941
          ||||||| || ||||| || ||||||||||||||||||||||||||||||||||||||||
Db     97 GTAGAGACGGGGTTTCACCTTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 38

Qy  12942 ACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 12978
           || ||||||||||||||||||||||||||||||||
Db     37 GCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 1


RESULT   3
HUMALCE221/c
LOCUS       HUMALCE221    103 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION   M87896
VERSION     M87896.1  GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       25 a     27 c     33 g     18 t
ORIGIN


  Query Match             0.5%;  Score 84.4;  DB 9;  Length 103;
  Best Local Similarity   89.2%;  Pred. No. 0.00069;
  Matches   91;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  12746 CTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGAT 12805
          |||||||||| |||  ||||||||||||||||||||||| ||||||||||||||| |||
Db    103 CTGGAGTGCAATGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAT 44

Qy  12806 TCTCCTGCCTCAGCCTCCCGAGTAGCTTGGACTACAGGCACA 12847
          |||||||||| ||| ||||| |||||| ||| ||||||||| |
Db     43 TCTCCTGCCTTAGCTTCCCGTGTAGCTGGGATTACAGGCAGA 2


RESULT   4
HSLDLI12
LOCUS       HSLDLI12      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low
            density lipoprotein.
ACCESSION   X05248
VERSION     X05248.1  GI:34334
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor;
            repetitive sequence.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     see X05249 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    complement(<1. .65)
                     /note="Alu repeat"
     intron          1. .108
                     /note="intron XII fragment"
BASE COUNT       21 a     38 c     20 g     29 t
ORIGIN


  Query Match             0.5%;  Score 83;  DB 10;  Length 108;
  Best Local Similarity   86.0%;  Pred. No. 0.0011;
  Matches   92;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  11214 TTGGCTCACTGCAACCTTTGCCTCTTGGGTTCAAGCAATTATCCTGCCTCAGCCTCCTTA 11273
          | | |||||  ||||||  ||||||  |||||||||  | |||  ||||||||||||||||||
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCTTA 61

Qy  11274 GTAGCAGGGTCTACAGGCATGTGCCACCACACCCGGCTGTTTTTGTA 11320
          |||||  |||  ||||  ||||||||||||||  ||||||||  |||||||
Db     62 GTAGCTGGGATTACAAGCATGTGCCACCACGCCCGGCTGATTTTGTA 108


RESULT   5
HUMALCE162/c
LOCUS       HUMALCE162    107 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION   M87924
VERSION     M87924.1  GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       28 a     30 c     35 g     14 t
ORIGIN


  Query Match             0.5%;  Score 82.2;  DB 9;  Length 107;
  Best Local Similarity   87.4%;  Pred. No. 0.0015;
  Matches   90;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  12712 TTTTTTTTGACTGAGTCTTGCTCTGTCTCACAGGCTGGAGTGCAGTGGTGCGATCTCGGC 12771
          ||||||  ||| ||||||  ||||||||  | ||||||||||||||||||  ||||||||||
Db    107 TTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGC 48

Qy  12772 TCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCC 12814
          |||||||||  |||  ||||||||||||||  |  |||||  |||||
Db     47 TCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTTCTGCC 5


RESULT   6
HSLDLRD1
LOCUS       HSLDLRD1      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION   X05249
VERSION     X05249.1  GI:34335
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05248 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     misc_feature    1. .108
                     /note="deletion junction region intron 12/ intron 15"
BASE COUNT       20 a     40 c     20 g     28 t
ORIGIN


  Query Match             0.4%;  Score 79.8;  DB 10;  Length 108;
  Best Local Similarity   84.1%;  Pred. No. 0.0034;
  Matches   90;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  13471 TTGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGACTCAGCCTCCCGA 13530
          | | ||||  ||||||||||||||||||||||  ||| |||||  ||||||||||||
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 61

Qy  13531 GTAGCTGGGATTACAGGCATGCATCACCATGCCTGGGTAATTTTTGT 13577
          ||||||||||||||||||      |||||  ||||||  ||||||||||
Db     62 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 108


RESULT   7
HSLDLRD2/c
LOCUS       HSLDLRD2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION   X05251
VERSION     X05251.1  GI:34336
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
```

```
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     20 c     40 g     20 t
ORIGIN


  Query Match             0.4%;  Score 79.8;  DB 10;  Length 108;
  Best Local Similarity   84.1%;  Pred. No. 0.0034;
  Matches   90;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  13471 TTGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGACTCAGCCTCCCGA 13530
          | | ||||   |||||||||||||||||||||||||   ||| ||||| ||||||||||||
Db    107 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 48

Qy  13531 GTAGCTGGGATTACAGGCATGCATCACCATGCCTGGGTAATTTTTGT 13577
          |||||||||||||||||     ||||| |||||| ||||||||||
Db     47 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   8
HSLDLRN2
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.4%;  Score 79.2;  DB 10;  Length 108;
  Best Local Similarity   83.3%;  Pred. No. 0.0042;
  Matches   90;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy   4290 ACAAAAAGTAGCCAGGCGTGGTGGTGCCCACCTGTAATCCCAGCTACTCAGGAGGCTGAA 4349
          ||||||| ||||||||||||||||      |||||||||||||||||| |||||||||
Db      1 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 60

Qy   4350 GCGGGAGGATCACTTGAACCTGGGAGGCAGAGATTGCAATGAGCTGAG 4397
          || |||| ||  ||||||||  |||||||||| ||||| ||||| |||
Db     61 GCAGGAGAATTGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAG 108


RESULT   9
HSU67804/c
LOCUS       HSU67804      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67804
VERSION     U67804.1  GI:2289918
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu3"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       26 a     38 c     26 g     18 t
ORIGIN


  Query Match             0.4%;  Score 77.8;  DB 11;  Length 108;
  Best Local Similarity   87.6%;  Pred. No. 0.0069;
  Matches   85;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  12882 GTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 12941
          | | ||| || ||||| || ||||||||||||||||||||||||||||||||||| |||||||
Db     97 GGAAAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTTGTGATCC 38

Qy  12942 ACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 12978
            || || | ||||| ||||||||||||||||||||
Db     37 TCCCGCTTTGGCCTTCCAAAGTGCTGGGATTACAGGC 1


RESULT  10
HUMUT8164A/c
LOCUS       HUMUT8164A     91 bp    DNA             STS       29-DEC-1994
DEFINITION  Human STS UT8164, 5' primer bind, sequence tagged site.
ACCESSION   L30244
VERSION     L30244.1  GI:605447
KEYWORDS    STS; PCR primer; STS sequence; microsatellite DNA; microsatellite
            marker; sequence tagged site; tetranucleotide repeat.
SOURCE      Homo sapiens DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
            Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
            Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
            White,R.
  TITLE     Genetic and physical mapping of simple sequence repeat containing
            sequence tagged sites from the human genome
  JOURNAL   Unpublished (1994)
COMMENT     Submitted by: Utah Center for Human Genome Research University of
```

```
          Utah, Dept. of Human Genetics
          2160 Eccles Institute of Human Genetics
          Salt Lake City, UT  84112
          e-mail: sts@corona.med.utah.edu
          Primer A:  AGAGGTTGCAGTGAACCAA
          Primer B:  TTTTTCCCCCTCTACTCACT
          End to Label: Primer B
          PCR Profile:
          Initial Denaturation: 94C 300sec
          Cycles     Denaturation    Annealing       Extension 5          94
          C  10 sec.   56 C  10 sec.   72 C  20 sec. 30         94 C  10 sec.
            52 C  10 sec.   72 C  20 sec. Mg++: 1.50 mM
          Gel: Acrylamide 7%, Formamide 32%, Urea 34%
          Alleles: 1.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     primer_bind     10. .28
                     /evidence=experimental
BASE COUNT       35 a     20 c     23 g     13 t
ORIGIN


  Query Match             0.4%;  Score 77.2;  DB 13;  Length 91;
  Best Local Similarity   91.1%;  Pred. No. 0.0088;
  Matches   82;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy  12704 TTTTTTTTTTTTTTTTGACTGAGTCTTGCTCTGTCTCACAGGCTGGAGTGCAGTGGTGCG 12763
          |||||||||||||||||| ||||||  ||||||| | |||||||||||||||||||| |
Db     91 TTTTTTTTTTTTTTTTGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGTG 32

Qy  12764 ATCTCGGCTCACTGCAACCTCTGCCTCCCG 12793
          |||| || ||||||||||||||||||||||
Db     31 ATCTTGGTTCACTGCAACCTCTGCCTCCCG 2


RESULT  11
HS8IC8R
LOCUS       HS8IC8R       103 bp    DNA             STS       05-SEP-1991
DEFINITION  Human sequence tagged site 8IC8R DNA from 19q13.
ACCESSION   X57789
VERSION     X57789.1  GI:23938
KEYWORDS    STS; myotonic dystrophy.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Aldridge,F.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
            Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE   2  (bases 1 to 103)
  AUTHORS   Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
            Davies,J., Johnson,K. and Markham,A.F.
  TITLE     Two sequence-tagged sites defining the ends of a 380 kb YAC clone
            from 19q13
  JOURNAL   Nucleic Acids Res. 19 (17), 4787 (1991)
  MEDLINE   91367697
COMMENT     See also X57788 for STS 8IC8L.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19q13"
                     /germline
                     /clone_lib="YAC library: ICI"
                     /clone="8IC8"
BASE COUNT       29 a     28 c     23 g     22 t      1 others
ORIGIN


  Query Match             0.4%;  Score 77.2;  DB 13;  Length 103;
  Best Local Similarity   85.9%;  Pred. No. 0.0086;
  Matches   85;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  11326 AGTAGAGATGGGGTTTCACCATGTTGGCCAGGATGATCTCGAACACCTGACCTCAAGTGA 11385
          ||||||||| ||||||||||||||||||||| || ||  |||| |||||||| ||||||
Db      2 AGTAGAGATAGGGTTTCACCATGTTGGCCAGGCTGGTCAAGAACTCCTGACCTTAAGTGA 61

Qy  11386 TCCACCCACCTCAGTCTCCCAAAGTGCTGGGATTACAGG 11424
          |||||||||||||   |||||||||||| |  |||| |||
Db     62 TCCACCCACCTCGACCTCCCAAAGTGCNGAAATTATAGG 100


RESULT  12
HSU67808/c
LOCUS       HSU67808      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67808
VERSION     U67808.1  GI:2289922
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu7"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       22 a     37 c     28 g     21 t
ORIGIN


  Query Match             0.4%;  Score 76.2;  DB 11;  Length 108;
  Best Local Similarity   86.6%;  Pred. No. 0.012;
  Matches   84;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  12882 GTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 12941
          | | ||| |||||||| || |||| |||||| ||||||| | |||||||||| |||||||
Db     97 GGAAAGACGGAGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTTGTGATCC 38

Qy  12942 ACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 12978
          ||| | || ||||||||||||||||||||||||||||
Db     37 ACCCGACTTGGCCTCCCAAAGTGCTGGGATTACAGGC 1


RESULT  13
HS8IC8R/c
LOCUS       HS8IC8R       103 bp    DNA             STS       05-SEP-1991
DEFINITION  Human sequence tagged site 8IC8R DNA from 19q13.
ACCESSION   X57789
VERSION     X57789.1  GI:23938
KEYWORDS    STS; myotonic dystrophy.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Aldridge,F.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
            Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE   2  (bases 1 to 103)
  AUTHORS   Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
            Davies,J., Johnson,K. and Markham,A.F.
  TITLE     Two sequence-tagged sites defining the ends of a 380 kb YAC clone
            from 19q13
  JOURNAL   Nucleic Acids Res. 19 (17), 4787 (1991)
  MEDLINE   91367697
COMMENT     See also X57788 for STS 8IC8L.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19q13"
                     /germline
                     /clone_lib="YAC library: ICI"
                     /clone="8IC8"
BASE COUNT       29 a     28 c     23 g     22 t      1 others
ORIGIN


  Query Match             0.4%;  Score 75;  DB 13;  Length 103;
  Best Local Similarity   84.0%;  Pred. No. 0.019;
  Matches   84;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   8133 GCCTGTAATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTT 8192
           |||| ||||  | |||||||||| |||  |||| |||  |||||||| ||||||||||||
Db    101 GCCTATAATTTCNGCACTTTGGGAGGTCGAGGTGGGTGGATCACTTAAGGTCAGGAGTTC 42

Qy   8193 GAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCCACT 8232
           |||||||||||||||||||||||||||||| |||| |||
Db     41 TTGACCAGCCTGGCCAACATGGTGAAACCCTATCTCTACT 2


RESULT  14
HUMALCE43/c
LOCUS       HUMALCE43     110 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION   M87900
VERSION     M87900.1  GI:174876
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       27 a     31 c     34 g     18 t
ORIGIN


  Query Match             0.4%;  Score 74.4;  DB 9;  Length 110;
  Best Local Similarity   80.6%;  Pred. No. 0.023;
  Matches   87;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy  12893 GTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGG 12952
          |||||| | |||||||||||| |||||| || || |||| ||||  |||| || |||| ||
Db    110 GTTTCGTCATGTTAGCCAGGCTGGTCTTGAACTACTGAGCTCGCAATCCTCCTGCCTTGG 51

Qy  12953 CCTCCCAAAGTGCTGGGATTACAGGCATGGGCCACCACGCCTGGCCCG 13000
          |||||||||||||| ||||||  |||  || |||||| |||| |||| |
Db     50 CCTCCCAAAGTGCCGGGATTGTAGGTGTGAGCCACCGCGCCCGGCCAG 3


RESULT  15
HUMD1D03M5/c
LOCUS       HUMD1D03M5    108 bp    mRNA            PRI       04-FEB-1999
DEFINITION  Human HepG2 partial cDNA, clone hmd1d03m5.
ACCESSION   D16965
VERSION     D16965.1  GI:598552
KEYWORDS    gene signature.
SOURCE      Homo sapiens Male cell_line:HepG2 cDNA to mRNA, clone_lib:Kiseru.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Matoba,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-1993) to the DDBJ/EMBL/GenBank databases. Ryo
            Matoba, Osaka University, Institute for Molecular and Cellular Bio;
            1-3, Yamada-oka, Suita, Osaka 565, Japan
            (E-mail:matoba@inherit.imcb.osaka-u.ac.jp,
            Tel:81-6-877-5111(ex.3314), Fax:81-6-877-1922)
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Matoba,R., Okubo,K., Hori,N., Fukushima,A. and Matsubara,K.
  TITLE     The addition of 5'-coding information to a 3'-directed cDNA library
            improves analysis of gene expression
  JOURNAL   Gene 146 (2), 199-207 (1994)
  MEDLINE   94357437
COMMENT     Submitted (21-Jul-1993) to DDBJ by:
            Ryo Matoba
            Molecular Microbiology and Genetics Lab.
            Research Institute of Innovative Technology for the Earth 9-2
            Kizugawadai Kizu-cyo,
            Soraku-gun, Kyoto
            Japan, 619-02
            Phone: 07747-5-2308
            Fax:   07747-5-2321.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="HepG2"
                     /clone_lib="Kiseru"
                     /sex="Male"
BASE COUNT       28 a     23 c     38 g     17 t      2 others
ORIGIN


  Query Match             0.4%;  Score 74.2;  DB 9;  Length 108;
  Best Local Similarity   85.3%;  Pred. No. 0.024;
  Matches   93;  Conservative    0;  Mismatches   15;  Indels    1;  Gaps    1;

Qy  13467 GATCTTGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGACTCAGCCTC 13526
          |||||||||||| |||||||||||||||| ||| ||||| || ||||||| |||||||||
Db    108 GATCTTGGCTCACTGCAACCTCTGCCTCCCGGGNTCAAGCGACTCTCCTGCCTCAGCCTC 49

Qy  13527 CCGAGTAGCTGGGATTACAGGCATGCATCACCATGCCTGGGTAATTTTT 13575
          | |||||||||||||||| ||||||  |||||  |||||  |  ||| |
Db     48 CTGAGTAGCTGGGATTACA-GCATGCGCCACCACNCCTGGCTTTTTTAT 1


Search completed: June 17, 2000, 10:40:36
Job time: 252120 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 11:43:03 ; Search time 939.94 Seconds
(without alignments)
4791.484 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTCACTTCTAACCA.....CCATCCAGCTTCTCACTCCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	71.8	0.4	108	1	X12095	Human biallelic po
	2	71.4	0.4	108	1	X01323	Allelic ladder, HU
c	3	68.6	0.4	108	1	X12095	Human biallelic po
c	4	65.6	0.4	100	1	T24892	Human gene signatu
c	5	65.2	0.4	103	1	T26213	Human gene signatu
c	6	64	0.4	108	1	T25009	Human gene signatu
	7	63.8	0.4	108	1	T26828	Human gene signatu
	8	60.2	0.3	100	1	T24892	Human gene signatu
c	9	58.6	0.3	103	1	T20927	Human gene signatu
c	10	58.4	0.3	87	1	T21566	Human gene signatu
c	11	57.4	0.3	85	1	Q95218	Simple tandem repe
	12	57	0.3	100	1	X12085	Human biallelic po
	13	57	0.3	100	1	X12086	Human biallelic po
c	14	56.8	0.3	93	1	T22572	Human gene signatu
c	15	56.4	0.3	82	1	T22836	Human gene signatu
c	16	56.4	0.3	91	1	T25854	Human gene signatu
	17	55.6	0.3	100	1	X12087	Human biallelic po
c	18	55.4	0.3	106	1	V11611	Homo sapiens adult
	19	54.8	0.3	103	1	T20927	Human gene signatu
c	20	54.2	0.3	93	1	T24259	Human gene signatu
	21	54	0.3	109	1	T23895	Human gene signatu
c	22	53.8	0.3	100	1	X12087	Human biallelic po
c	23	53.8	0.3	100	1	X12085	Human biallelic po
c	24	53.8	0.3	100	1	X12086	Human biallelic po
	25	52.6	0.3	95	1	T23131	Human gene signatu
c	26	52.6	0.3	97	1	T26728	Human gene signatu
	27	52.4	0.3	87	1	T21566	Human gene signatu
c	28	52.2	0.3	107	1	T20373	Human gene signatu
c	29	52.2	0.3	109	1	T23895	Human gene signatu
	30	52	0.3	69	1	Q29016	Probe to internal
c	31	52	0.3	69	1	Q29016	Probe to internal
	32	52	0.3	110	1	T26288	Human gene signatu
c	33	51.6	0.3	99	1	T23728	Human gene signatu
	34	51	0.3	70	1	N60231	Normal chromosome
	35	51	0.3	91	1	T25854	Human gene signatu
	36	51	0.3	99	1	T20931	Human gene signatu
c	37	50.8	0.3	85	1	T25730	Human gene signatu
c	38	50.6	0.3	110	1	T26288	Human gene signatu
c	39	50.4	0.3	92	1	V11595	Homo sapiens adult
c	40	50.4	0.3	92	1	V61480	Human secreted pro
c	41	50	0.3	108	1	T26828	Human gene signatu
c	42	49.6	0.3	93	1	T25688	Human gene signatu
	43	49.4	0.3	97	1	T26728	Human gene signatu
	44	49	0.3	65	1	T25588	Human gene signatu
	45	48.4	0.3	75	1	T22841	Human gene signatu

ALIGNMENTS

RESULT 1
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.4%; Score 71.8; DB 1; Length 108;
Best Local Similarity 84.9%; Pred. No. 0.025;
Matches 79; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

```
Qy  13575 TGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGACCAGGCTGGTCTCAAACTCCTGAC 13634
          ||| ||||| ||||||||| ||||| | :||||| ||||| ||||||| ||||||||||
Db      1 TGTCTTTTTTGTAGAGATGAGGTTTTCCTRTGTTGGCCAGGATGGTCTCGAACTCCTGAC 60

Qy  13635 CTCAAGTGATCCACCTGCCTTAGCCTCCCAAAA 13667
          ||||||||||| ||||||| ||||||||||||
Db     61 TTCAAGTGATCCGTCTGCCTTGGCCTCCCAAAA 93
```

```
RESULT   2
X01323
ID   X01323 standard; DNA; 108 BP.
AC   X01323;
DT   14-APR-1999  (first entry)
DE   Allelic ladder, HUMFIBRA/FGA allele 27.
KW   Allelic ladder; human; HUMVWFA31/A; HUMTHO1; D8S1179; HUMFIBRA/FGA; AMG;
KW   D21S11; D18S51; forensic testing; DNA profiling; amelogenin sex test; ss.
OS   Homo sapiens.
PN   EP-887426-A2.
PD   30-DEC-1998.
PF   29-JUN-1998; 305120.
PR   28-JUN-1997; GB-013597.
PA   (UKHO-) UK SEC STATE HOME DEPT.
PI   Arnold CD, Barber MD, Burke T, Gill P, Gillbard SM,
PI   Griffiths RA, Haywood MD, Johnson PE, Smith CD, Urquhart AJ;
DR   WPI; 99-047890/05.
PT   New alleles and allelic ladder mixtures - useful as a control sample
PT   for DNA profiling in forensic environments
PS   Claim 1; Page 39; 52pp; English.
CC   This sequence represents an allelic ladder, used in the allelic ladder
CC   mixture of the invention. The allelic ladder mixture comprises at least
CC   one of the following allelic ladders: (i) at least one allele for locus
CC   HUMVWFA31/A; (ii) one allele for locus HUMTHO1; (iii) at least one allele
CC   for locus D8S1179; (iv) at least one allele for locus HUMFIBRA/FGA;
CC   (v) at least one allele for locus D21S11; and (vi) an allele for locus
CC   D18S51. The alleles or allelic ladders are useful in forensic tests for
CC   comparison with a sample DNA profile, when the profile is based on
CC   analysis of at least one loci of HUMVWFA31/A, HUMTHO1, D8S1179,
CC   HUMFIBRA/FGA, D21S11, D18S51 or AMG (amelogenin sex test). DNA profiling
CC   is useful in anthropological, paternity, maternity, crime detection and
CC   other forensic environments. The new allelic ladders have an improved
CC   range and coverage of DNA, and include a number of rare alleles which
CC   offers improved identification of an unknown sample.
SQ   Sequence   108 BP;    0 A;    31 C;    0 G;    77 T;


  Query Match             0.4%;  Score 71.4;  DB 1;  Length 108;
  Best Local Similarity   83.5%;  Pred. No. 0.028;
  Matches   81;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   5756 TCCTTTCTTCTCTTTCTCTCTTTCTTTCTTTCTCTCTTTCTCTTTCTTTCTTTCTCTCTC 5815
          | |||||||   ||| ||||||||||||||||| ||||||| | | | | | | ||| ||
Db      2 TTCTTTCTTTCTTTTTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTC 61

Qy   5816 TCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 5852
          | |||||||||| |||||||||||||||||||||||||
Db     62 TTTCTTTCTTTCCTTCTTTCTTTCTTTCTTTCTTTCT 98


RESULT   3
X12095/c
ID   X12095 standard; DNA; 108 BP.
AC   X12095;
DT   30-MAR-1999  (first entry)
DE   Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW   Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW   detection; phenotypic typing; characteristic; infection; hereditary;
KW   autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW   treatment; marker; ss.
OS   Homo sapiens.
PN   WO9820165-A2.
PD   14-MAY-1998.
PF   05-NOV-1997; U20313.
PR   06-NOV-1996; US-030455.
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI   Hudson T, Lander ES, Wang D;
DR   WPI; 98-286974/25.
PT   New isolated nucleic acid segments from the human genome - used for
PT   determining polymorphic forms for use in e.g. forensics, paternity
PT   testing or phenotypic typing for disease
PS   Claim 1; Page 219; 310pp; English.
CC   X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC   markers which have been isolated using the primers represented in
CC   X09121-X10268. The base occupying the polymorphic site is indicated by
CC   the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC   methods for determining polymorphic forms in an individual for use in
CC   e.g. forensics, paternity testing or for phenotypic typing for diseases
CC   such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC   muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC   hypercholesterolemia, polycystic kidney disease, hereditary
CC   spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC   haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC   syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC   autoimmune diseases, inflammation, cancer, diseases of the nervous
CC   system, infection by pathogenic microorganisms, and characteristics such
CC   as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC   endurance, fertility, and susceptibility or receptivity to particular
CC   drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC   segments can also be used to produce medicaments for the treatment or
CC   prophylaxis of such diseases.
SQ   Sequence   108 BP;    19 A;    23 C;    28 G;    37 T;


  Query Match             0.4%;  Score 68.6;  DB 1;  Length 108;
  Best Local Similarity   83.8%;  Pred. No. 0.068;
  Matches   88;  Conservative    1;  Mismatches   15;  Indels    1;  Gaps    1;

Qy   4189 TAATCCCAGCAC-TTTGGGAGGCCGAGGCAGGAGGATCACATGAGGTCAGAAATTCAAGA 4247
          |||||||||||| |||||||||| ||||||  ||||||| ||| ||||| | ||| |||
Db    105 TAATCCCAGCACTTTTGGGAGGCCAAGGCAGACGGATCACTTGAAGTCAGGAGTTCGAGA 46

Qy   4248 CCAGCCTGGCCAAAATGGTGAAACCTCATCTCTACTAAAAATACA 4292
          ||| ||||||||| |: |  |||||||||||||| ||||| |||
Db     45 CCATCCTGGCCAACAYAGGAAAACCTCATCTCTACAAAAAAGACA 1


RESULT   4
T24892/c
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;
```

```
  Query Match             0.4%;  Score 65.6;  DB 1;  Length 100;
  Best Local Similarity   77.8%;  Pred. No. 0.18;
  Matches   77;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy  12707 TTTTTTTTTTTTTGACTGAGTCTTGCTCTGTCTCACAGGCTGGAGTGCAGTGGTGCGATC 12766
          ||| ||| |||   || |||| |  ||||||| | ||||| ||||||||  ||||| |||
Db    100 TTTGTTTGTTTCAAACAGAGTGTCACTCTGTCACCCAGGCNGGAGTGCAANGGTGCAATC 41

Qy  12767 TCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGAT 12805
          || ||| | |||||  |||||||||| |||||||| |||
Db     40 TCAGCTNATTGCAAATTCTGCCTCCCAGGTTCAAGCGAT 2


RESULT   5
T26213/c
ID   T26213 standard; cDNA to mRNA; 103 BP.
AC   T26213;
DT   13-NOV-1996  (first entry)
DE   Human gene signature HUMGS08452.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2029; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    33 A;    21 C;    25 G;    23 T;


  Query Match             0.4%;  Score 65.2;  DB 1;  Length 103;
  Best Local Similarity   77.5%;  Pred. No. 0.2;
  Matches   79;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy   5856 TTTCTTTTTCTTTCTGACAGGGTCTTGCTCTATTGCCTAGGCTGGAGTGCAGTGGTGCAA 5915
          ||| ||||||||   |||| | |||| |||| | ||| ||||||||||||||||||| |
Db    102 TTTTTTTTTCTTAAAGACATGTTCTTACTCTGTGGCCCAGGCTGGAGTGCAGTGGTGCCA 43

Qy   5916 TCTCAGCTCACTGCAGCCTTGAACTCCAGGGCTCAAGCAATC 5957
          ||  ||||||||| | |    |||||| || ||||||  |||
Db     42 TCATAGCTCACTGTAACACCAAACTCCTGGACTCAAGTGATC 1


RESULT   6
T25009/c
ID   T25009 standard; cDNA to mRNA; 108 BP.
AC   T25009;
DT   07-NOV-1996  (first entry)
DE   Human gene signature HUMGS07131.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1748; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    34 A;    31 C;    26 G;    15 T;


  Query Match             0.4%;  Score 64;  DB 1;  Length 108;
  Best Local Similarity   74.5%;  Pred. No. 0.3;
  Matches   79;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy  15026 TTTGTTTGTTTGTTTTTTGAGATAGAGTCTTCCTCTGTCATTCAGGCTGGAGTGCAGTGG 15085
          ||  | |  ||||| |||   | || ||||| |||||||| |||||||||| | ||||||
Db    107 TTGNTGTTGTTGTTGTTTTCAACAGGGTCTTGCTCTGTCACTCAGGCTGGAATNCAGTGG 48

Qy  15086 CATGATCTCAGCTCACTGCAGCCTCCGCCTCCCGGGTTCAAGAGAT 15131
          | ||| |   |||||||||||||  |||||  ||| ||| | |||
Db     47 CGTGACCATGGCTCACTGCAGCCTTGGCCTCATGGGCTCAGGCGAT 2


RESULT   7
T26828
ID   T26828 standard; cDNA to mRNA; 108 BP.
AC   T26828;
DT   14-NOV-1996  (first entry)
DE   Human gene signature HUMGS09078.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2182; 2245pp; Japanese.
```

```
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    18 A;    33 C;    23 G;    28 T;


  Query Match             0.4%;  Score 63.8;  DB 1;  Length 108;
  Best Local Similarity   88.3%;  Pred. No. 0.32;
  Matches   68;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  12921 GATCTCCTGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCAT 12980
          ||||||||||||||||| |||| || |  |||||||||| |||||||| ||||||||||
Db      1 GATCTCCTGACCTCGTCATCCGCCCGTNTCGGCCTCCCATAGTGCTGGGNTTACAGGCAT 60

Qy  12981 GGGCCACCACGCCTGGC 12997
          | ||||||||||| |||
Db     61 GAGCCACCACGCCCGGC 77


RESULT   8
T24892
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;


  Query Match             0.3%;  Score 60.2;  DB 1;  Length 100;
  Best Local Similarity   74.0%;  Pred. No. 1;
  Matches   74;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy   4357 GATCACTTGAACCTGGGAGGCAGAGATTGCAATGAGCTGAGATCATGCCACTGCACTCCA 4416
          |||| ||||||||||||||||||  ||||||| |||||||||    ||  ||||||||
Db      1 GATCGCTTGAACCTGGGAGGCAGAATTTGCAATNAGCTGAGATTGCACCNTTGCACTCCN 60

Qy   4417 GCCTGGGCGACAGTGAGACTCCATCTCAAAATAATAATAA 4456
          ||||||| ||||| | ||| |  | |   ||  | |  ||
Db     61 GCCTGGGTGACAGAGTGACACTCTGTTTGAAACAAACAAA 100


RESULT   9
T20927/c
ID   T20927 standard; cDNA to mRNA; 103 BP.
AC   T20927;
DT   24-JUL-1996  (first entry)
DE   Human gene signature HUMGS02180.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 758-759; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    22 A;    27 C;    21 G;    31 T;


  Query Match             0.3%;  Score 58.6;  DB 1;  Length 103;
  Best Local Similarity   73.7%;  Pred. No. 1.7;
  Matches   73;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy   4262 ATGGTGAAACCTCATCTCTACTAAAAATACAAAAAGTAGCCAGGCGTGGTGGTGCCCACC 4321
          |||| ||||     || ||||| ||||||| |||| |||  | | ||||||  | ||||
Db     99 ATGGAGAAATACTGTCCCTACTNAAAATACNAAAATCAGCTGGACATGGTGGCACACACC 40

Qy   4322 TGTAATCCCAGCTACTCAGGAGGCTGAAGCGGGAGGATC 4360
          ||||  | ||||||||  |||||  |||| |||||||||
Db     39 TGTAGCCACAGCTACTTGGGAGGTGGAAGTGGGAGGATC 1


RESULT  10
T21566/c
ID   T21566 standard; cDNA to mRNA; 87 BP.
AC   T21566;
DT   03-AUG-1996  (first entry)
DE   Human gene signature HUMGS02944.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
```

KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 914; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 87 BP; 35 A; 21 C; 16 G; 13 T;

Query Match 0.3%; Score 58.4; DB 1; Length 87;
Best Local Similarity 79.1%; Pred. No. 1.8;
Matches 68; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 13560 TGCCTGGGTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGACCAGGCTGG 13619
Db 86 TTCTTGGCTNATTTCTGTATTTTTTGTAAAGATGGGGTTTCGCCATGTTCCTCGGGCTGG 27

Qy 13620 TCTCAAACTCCTGACCTCAAGTGATC 13645
Db 26 TTTTAAACTCCTGGGNTCAAGCGATC 1

RESULT 11
Q95218/c
ID Q95218 standard; DNA; 85 BP.
AC Q95218;
DT 08-FEB-1996 (first entry)
DE Simple tandem repeat (STR) corresponding to the 2nd part of wg3b10.
KW Simple tandem repeat; STR; wg3b10; treatment; genetic; diagnosis;
KW characterisation; mapping; linkage studies; analysis; alleles;
KW second part; ss.
OS Synthetic.
PN WO9517522-A2.
PD 29-JUN-1995.
PF 21-DEC-1994; G02789.
PR 21-DEC-1993; GB-026052.
PA (UYLE-) UNIV LEICESTER.
PI Armour J, Jeffreys AJ;
DR WPI; 95-240682/31.
PT Identifying simple tandem repeat loci in DNA - by screening DNA
PT library to enrich for fragments contg. the repeats before cloning
PT and rescreening, also simple tandem repeats for treatment or
PT diagnosis
PS Claim 26; Page 29; 51pp; English.
CC Q95218 is a simple tandem repeat (STR) corresponding to the 2nd part
CC of wg3b10. The STR can be used for treatment and diagnosis in human
CC and veterinary medicine, partic. for genetic characterisation,
CC mapping, linkage studies and analysis/diagnosis of acquired disease
CC alleles.
SQ Sequence 85 BP; 58 A; 1 C; 24 G; 2 T;

Query Match 0.3%; Score 57.4; DB 1; Length 85;
Best Local Similarity 80.7%; Pred. No. 2.5;
Matches 67; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 5772 TCTCTTTCTTTCTTTCTCTCTTTCTCTTTCTTTCTTTCTCTCTCTCTTTCTTTCTTTC 5831
Db 83 TCCCTTTCATTCCTTTCCTTTTTTTCTTTTCTTTCTTTCTTTCCTCTCTTTCTTTCTTTC 24

Qy 5832 TTTCTTTCTTTCTTTCTTTCTTT 5854
Db 23 TTTCTTTCTTTCTTTCTTTCTTT 1

RESULT 12
X12085
ID X12085 standard; DNA; 100 BP.
AC X12085;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment EST98276c.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 218; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 100 BP; 22 A; 25 C; 22 G; 30 T;

Query Match 0.3%; Score 57; DB 1; Length 100;
Best Local Similarity 72.7%; Pred. No. 2.8;
Matches 72; Conservative 1; Mismatches 26; Indels 0; Gaps 0;

Qy 11336 GGGTTTCACCATGTTGGCCAGGATGATCTCGAACACCTGACCTCAAGTGATCCACCCACC 11395
Db 1 GAGTCTTGCTATGTTTCCCAGGATGGTCTTGAGCTCCTGGTTTCAAACAATCCTCCTTCC 60

Qy 11396 TCAGTCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCAC 11434
Db 61 TAAGCCTCCYAAAGTGCCAGGATTATAGGTGTGAGTCAC 99

```
RESULT  13
X12086
ID   X12086 standard; DNA; 100 BP.
AC   X12086;
DT   30-MAR-1999  (first entry)
DE   Human biallelic polymorphic DNA fragment EST98276b.
KW   Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW   detection; phenotypic typing; characteristic; infection; hereditary;
KW   autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW   treatment; marker; ss.
OS   Homo sapiens.
PN   WO9820165-A2.
PD   14-MAY-1998.
PF   05-NOV-1997; U20313.
PR   06-NOV-1996; US-030455.
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PI   Hudson T, Lander ES, Wang D;
DR   WPI; 98-286974/25.
PT   New isolated nucleic acid segments from the human genome - used for
PT   determining polymorphic forms for use in e.g. forensics, paternity
PT   testing or phenotypic typing for disease
PS   Claim 1; Page 219; 310pp; English.
CC   X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC   markers which have been isolated using the primers represented in
CC   X09121-X10268. The base occupying the polymorphic site is indicated by
CC   the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC   methods for determining polymorphic forms in an individual for use in
CC   e.g. forensics, paternity testing or for phenotypic typing for diseases
CC   such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC   muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC   hypercholesterolemia, polycystic kidney disease, hereditary
CC   spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC   haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC   syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC   autoimmune diseases, inflammation, cancer, diseases of the nervous
CC   system, infection by pathogenic microorganisms, and characteristics such
CC   as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC   endurance, fertility, and susceptibility or receptivity to particular
CC   drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC   segments can also be used to produce medicaments for the treatment or
CC   prophylaxis of such diseases.
SQ   Sequence   100 BP;    21 A;    25 C;    22 G;    31 T;


  Query Match             0.3%;  Score 57;  DB 1;  Length 100;
  Best Local Similarity   72.7%;  Pred. No. 2.8;
  Matches   72;  Conservative    1;  Mismatches   26;  Indels    0;  Gaps    0;

Qy  11336 GGGTTTCACCATGTTGGCCAGGATGATCTCGAACACCTGACCTCAAGTGATCCACCCACC 11395
          | || |  | |||||  |||||||| ||| || | ||||   ||||   |||| ||  ||
Db      1 GAGTCTTGCTATGTTTCCCAGGATGGTCTTGAGCTCCTGGTTTCAAACAATCCTCCTTCC 60

Qy  11396 TCAGTCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCAC 11434
          |:|| |||| |||||||  |||||| |||||||| |||
Db     61 TMAGCCTCCTAAAGTGCCAGGATTATAGGTGTGAGTCAC 99


RESULT  14
T22572/c
ID   T22572 standard; cDNA to mRNA; 93 BP.
AC   T22572;
DT   01-OCT-1996  (first entry)
DE   Human gene signature HUMGS04188.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1159; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   93 BP;    22 A;    23 C;    24 G;    22 T;


  Query Match             0.3%;  Score 56.8;  DB 1;  Length 93;
  Best Local Similarity   77.9%;  Pred. No. 3;
  Matches   67;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy   5872 ACAGGGTCTTGCTCTATTGCCTAGGCTGGAGTGCAGTGGTGCAATCTCAGCTCACTGCAG 5931
          ||||||||||||   ||  || | ||||||||||||||||  | |  ||||||||||||
Db     86 ACAGGGTCTTGCATNATAACCAANGCTGGAGTGCAGTGGTGTGACCATAGCTCACTGCAG 27

Qy   5932 CCTTGAACTCCAGGGCTCAAGCAATC 5957
          |||  | ||||  ||||||||  |||
Db     26 CCTCAACCTCCTCGGCTCAAGGGATC 1


RESULT  15
T22836/c
ID   T22836 standard; cDNA to mRNA; 82 BP.
AC   T22836;
DT   29-AUG-1996  (first entry)
DE   Human gene signature HUMGS04510.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1224; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
```

CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 82 BP; 26 A; 25 C; 14 G; 17 T;

Query Match 0.3%; Score 56.4; DB 1; Length 82;
Best Local Similarity 80.5%; Pred. No. 3.4;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy   4542 TTTAAGTTCCTGGGTACAAGTACAGGATGTGCAGGTTTGTTACATAGGTAAACGTGTGCC 4601
          |||||||||  |||||||  | ||||||||||||||||||||   ||||||| |||||||
Db     82 TTTAAGTTCTGGGGTACACATGCAGGATGTGCAGGTTTGTTATGCAGGTAAATGTGTGCC 23

Qy   4602 ATGGTGATTTGCTGCACCTATC 4623
          | || | |||||| | |  |||
Db     22 ACGGAGGTTTGCTCCTCAGATC 1
```

Search completed: June 17, 2000, 11:43:23
Job time: 254714 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 02:34:25 ; Search time 13753.1 Seconds
(without alignments)
5305.117 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTCACTTCTAACCA.....CCATCCAGCTTCTCACTCCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query

	No.	Score	Match	Length	DB	ID	Description
	1	93	0.5	109	30	AA243009	AA243009 zr25h02.s
	2	91.6	0.5	106	37	AA703692	AA703692 ag81a10.r
	3	91	0.5	107	35	AA565533	AA565533 nk42b11.s
c	4	90.2	0.5	103	84	B48914	B48914 RPCI11-4A12
c	5	90	0.5	110	106	AQ386882	AQ386882 RPCI11-13
	6	88.6	0.5	103	108	AQ535244	AQ535244 RPCI-11-3
	7	87.6	0.5	110	106	AQ386882	AQ386882 RPCI11-13
c	8	87.4	0.5	106	108	AQ544957	AQ544957 CITBI-E1-
	9	86.8	0.5	110	39	AA897366	AA897366 am06h02.s
	10	86.6	0.5	103	38	AA807640	AA807640 nx08b05.s
	11	86.6	0.5	109	84	B17434	B17434 345K2.TVB C
c	12	86.6	0.5	109	84	B17434	B17434 345K2.TVB C
	13	86.6	0.5	110	30	AA244245	AA244245 nc07a04.s
	14	86.6	0.5	110	94	AQ003188	AQ003188 RPCI11-1D
c	15	86.2	0.5	107	35	AA565533	AA565533 nk42b11.s
c	16	86.2	0.5	108	84	B65160	B65160 CIT-HSP-201
c	17	86	0.5	105	109	AQ637292	AQ637292 RPCI-11-4
c	18	85.4	0.5	103	94	AQ028649	AQ028649 CIT-HSP-2
c	19	85.4	0.5	103	108	AQ535244	AQ535244 RPCI-11-3
	20	85	0.5	106	30	AA250812	AA250812 zs06a05.s
c	21	84.4	0.5	102	36	AA654562	AA654562 nt75f10.s
c	22	83.6	0.5	106	105	AQ264176	AQ264176 CITBI-E1-
	23	83.4	0.5	109	105	AQ265749	AQ265749 CITBI-E1-
c	24	83	0.5	107	35	AA583252	AA583252 nn41e04.s
	25	83	0.5	109	94	AQ029690	AQ029690 RPCI11-41
	26	82	0.5	103	108	AQ582186	AQ582186 RPCI-11-4
	27	82	0.5	106	63	AI991750	AI991750 wt48e01.x
c	28	82	0.5	106	63	AI991750	AI991750 wt48e01.x
	29	81.6	0.5	107	24	H67040	H67040 yu68c01.r1
	30	81.6	0.5	108	35	AA594869	AA594869 no21e02.s
c	31	81.2	0.5	102	84	B48088	B48088 RPCI11-4N6.
c	32	81.2	0.5	104	108	AQ544583	AQ544583 CITBI-E1-
	33	81	0.4	105	28	AA078003	AA078003 7H12D08 C
	34	80.8	0.4	107	106	AQ412658	AQ412658 RPCI-11-2
	35	80.8	0.4	108	84	B32951	B32951 HS-1016-A1-
c	36	80.8	0.4	109	103	AQ200347	AQ200347 RPCI11-43
c	37	80.6	0.4	103	35	AA570438	AA570438 nk63g02.s
	38	80.6	0.4	103	108	AQ534922	AQ534922 RPCI-11-3
	39	80.6	0.4	108	84	B15423	B15423 345B10.TV C
	40	80.4	0.4	106	38	AA812141	AA812141 ob48h02.s
	41	80.4	0.4	109	24	N25299	N25299 yw52c09.s1
c	42	80.2	0.4	106	30	AA250812	AA250812 zs06a05.s
	43	80.2	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
c	44	80.2	0.4	109	105	AQ265749	AQ265749 CITBI-E1-
	45	80	0.4	97	39	AA837701	AA837701 oe06c02.s

ALIGNMENTS

```
RESULT   1
AA243009
LOCUS       AA243009      109 bp    mRNA            EST       11-MAR-1998
DEFINITION  zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
            cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
            element;contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION   AA243009
VERSION     AA243009.1  GI:1873869
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Dec 3, 1996 this sequence version replaced gi:1126869.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1127   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 102.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="GDB:5426481"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:664467"
                     /clone_lib="Stratagene NT2 neuronal precursor 937230"
                     /tissue_type="neuroepithelial cells"
                     /dev_stage="Ntera-2 neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Uninduced, exponentially growing neuroepithelial
                     cells (Ntera-2/c1.D1). Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     30 c     30 g     30 t
ORIGIN


  Query Match             0.5%;  Score 93;  DB 30;  Length 109;
  Best Local Similarity   90.8%;  Pred. No. 0.062;
  Matches   99;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  12873 GTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACC 12932
          ||||||||||||||| || ||||| |||||||||||||||||||| |||||||  |||
Db      1 GTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTTGATCTCCCTACC 60

Qy  12933 TCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 12981
          ||||||||| ||  ||||||||||||||||||||||||||||||||| ||
Db     61 TCGTGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 109


RESULT   2
AA703692
LOCUS       AA703692      106 bp    mRNA            EST       24-DEC-1997
DEFINITION  ag81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:1140858 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA703692
VERSION     AA703692.1  GI:2713610
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1397630.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
           Seq primer: -28m13 rev1 ET from Amersham
           High quality sequence stop: 53.
FEATURES             Location/Qualifiers
     source          1. .106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1140858"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a      29 c      29 g      29 t
ORIGIN


  Query Match             0.5%;  Score 91.6;  DB 37;  Length 106;
  Best Local Similarity   91.5%;  Pred. No. 0.091;
  Matches   97;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  12876 TTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 12935
          ||||||||||||| |  ||||| |||||||||||||||||||||||||||||||||||||
Db      1 TTTTTAGTAGAGACGAGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 60

Qy  12936 TGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 12981
          |||||  || ||||| ||||||||||||||||||||||||||| ||
Db     61 TGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG 106


RESULT   3
AA565533
LOCUS       AA565533      107 bp    mRNA            EST       08-SEP-1997
DEFINITION  nk42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA565533
VERSION     AA565533.1  GI:2337172
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393355.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1661   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1016157"
                     /clone_lib="NCI_CGAP_GC2"
                     /tissue_type="germ cell tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Bulk
                     germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCACGAG 3'
                     3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                     Average insert size:  1.2 kb."
BASE COUNT       22 a      34 c      26 g      25 t
ORIGIN


  Query Match             0.5%;  Score 91;  DB 35;  Length 107;
  Best Local Similarity   90.7%;  Pred. No. 0.11;
  Matches   97;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  12746 CTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGAT 12805
          |||||||||||||  | ||||| |||||||||| |||||||||| ||||||||||
Db      1 CTGGAGTGCAGTGGCTCAATCTCAGCTCACTGCAAGCTCTGCCTCCCAGGTTCAAGTGAT 60

Qy  12806 TCTCCTGCCTCAGCCTCCCGAGTAGCTTGGACTACAGGCACACACCA 12852
          |||| |||||||||||| |||||||| ||| ||||||||||||||
Db     61 TCTCGTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACACACCA 107


RESULT   4
B48914/c
LOCUS       B48914        103 bp    DNA             GSS       08-APR-1999
DEFINITION  RPCI11-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
            genomic survey sequence.
ACCESSION   B48914
VERSION     B48914.1  GI:2601151
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7501163"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-4A12"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       30 a      28 c      30 g      15 t
ORIGIN


  Query Match             0.5%;  Score 90.2;  DB 84;  Length 103;
```

```
  Best Local Similarity   92.2%;  Pred. No. 0.13;
  Matches   95;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy  12872 TGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGAC 12931
          ||||||||||||||||| || ||||| |||| |||||| ||||||||||||||||||||
Db    103 TGTATTTTAGTAGAGACGGGGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGAC 44

Qy  12932 CTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTAC 12974
          |||||||||| || ||||||||||||||||||||||| ||||
Db     43 CTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGCTTAC 1


RESULT   5
AQ386882/c
LOCUS       AQ386882      110 bp    DNA             GSS       21-MAY-1999
DEFINITION  RPCI11-134I4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-134I4,
            genomic survey sequence.
ACCESSION   AQ386882
VERSION     AQ386882.1  GI:4357905
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-134I4.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="GDB:7551267"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-134I4"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       26 a     26 c     38 g     20 t
ORIGIN


  Query Match             0.5%;  Score 90;  DB 106;  Length 110;
  Best Local Similarity   90.6%;  Pred. No. 0.14;
  Matches   96;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  13594 GGGTTTCACCATGTTGACCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCCACCTGCC 13653
          |||||||||||||||| ||||||||||||  |||||||||||||||| ||||||||||||
Db    110 GGGTTTCACCATGTTGTCCAGGCTGGTCTTGAACTCCTGACCTCAAGCGATCCACCTGCC 51

Qy  13654 TTAGCCTCCCAAAATGCTGGGACTACAGGCGTGAGCCACTGCACCC 13699
          | |||||||||||| | || ||| |||||||||||||||||||| |||
Db     50 TCAGCCTCCCAAAGTACTTGGATTACAGGCGTGAGCCACTGCTCCC 5


RESULT   6
AQ535244
LOCUS       AQ535244      103 bp    DNA             GSS       18-MAY-1999
DEFINITION  RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-317H22, genomic survey sequence.
ACCESSION   AQ535244
VERSION     AQ535244.1  GI:4846934
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7621533"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-317H22"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       31 a     27 c     27 g     18 t
ORIGIN


  Query Match             0.5%;  Score 88.6;  DB 108;  Length 103;
  Best Local Similarity   91.3%;  Pred. No. 0.2;
  Matches   94;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   8143 CCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCAGCC 8202
           |||||||||||| ||||  || ||||||||||||||||||||||||||||| |||||||||
Db      1 CCAGCACTTTGGGAGGCCAAGACGGGCAGATCACTTGAGGTCAGGAGTTCGAGACCAGCC 60

Qy   8203 TGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||||||||||||| |||||  ||| ||||||||||
Db     61 TGGCCAACATGGTGAAACCCCGTCTCTGCTATAAATACAAAAA 103


RESULT   7
AQ386882
LOCUS       AQ386882      110 bp    DNA             GSS       21-MAY-1999
DEFINITION  RPCI11-134I4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-134I4,
            genomic survey sequence.
ACCESSION   AQ386882
VERSION     AQ386882.1  GI:4357905
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
```

```
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: RPCI11-134I4.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .110
                      /organism="Homo sapiens"
                      /db_xref="GDB:7551267"
                      /db_xref="taxon:9606"
                      /clone="RPCI-11-134I4"
                      /clone_lib="RPCI-11"
                      /sex="Male"
                      /cell_type="Lymphocytes"
                      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                      RPCI11 Human Male BAC Library"
BASE COUNT       26 a     26 c     38 g     20 t
ORIGIN


  Query Match             0.5%;  Score 87.6;  DB 106;  Length 110;
  Best Local Similarity   87.3%;  Pred. No. 0.26;
  Matches   96;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    8113 GGCCGGGCCTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGA 8172
           |||||||    |||||||||||||||||||| || ||||||| |||||||||| ||  ||
Db       1 GGCCGGGAGCAGTGGCTCACGCCTGTAATCCAAGTACTTTGGGAGGCTGAGGCAGGTGGA 60

Qy    8173 TCACTTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCC 8222
           || ||||||||||||||  |||||||||||| ||||||||||||||||
Db      61 TCGCTTGAGGTCAGGAGTTCAAGACCAGCCTGGACAACATGGTGAAACCC 110


RESULT   8
AQ544957/c
LOCUS       AQ544957      106 bp    DNA             GSS       28-MAY-1999
DEFINITION  CITBI-E1-2629N2.TF CITBI-E1 Homo sapiens genomic clone 2629N2,
            genomic survey sequence.
ACCESSION   AQ544957
VERSION     AQ544957.1  GI:4903683
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .106
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="2629N2"
                      /clone_lib="CITBI-E1"
                      /sex="male"
                      /cell_type="sperm"
                      /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                      CalTech Human BAC Library D"
BASE COUNT       26 a     26 c     36 g     18 t
ORIGIN


  Query Match             0.5%;  Score 87.4;  DB 108;  Length 106;
  Best Local Similarity   89.5%;  Pred. No. 0.28;
  Matches   94;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  12894 TTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGGC 12953
          |||| |||||||||||||| ||||||| ||||||||||||  |||||  || ||||||||
Db    106 TTTCACCGTGTTAGCCAGAATGGTCTTGATCTCCTGACCTAATGATCTGCCCGCCTCGGC 47

Qy  12954 CTCCCAAAGTGCTGGGATTACAGGCATGGGCCACCACGCCTGGCC 12998
          |||||||||||||||||||||||| |||||| ||||||| ||||
Db     46 CTCCCAAAGTGCTGGGATTACAGGCTTGGGCCTCCACGCCCGGCC 2


RESULT   9
AA897366
LOCUS       AA897366      110 bp    mRNA            EST       04-JAN-1999
DEFINITION  am06h02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1466067 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA897366
VERSION     AA897366.1  GI:3033986
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2150764.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 834   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 63.
FEATURES              Location/Qualifiers
     source           1. .110
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1466067"
                      /clone_lib="Soares_NFL_T_GBC_S1"
                      /lab_host="DH10B"
                      /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                      a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                      Equal amounts of plasmid DNA from three normalized
                      libraries (fetal lung NbHL19W, testis NHT, and B-cell
```

NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN

Query Match 0.5%; Score 86.8; DB 39; Length 110;
Best Local Similarity 88.7%; Pred. No. 0.32;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 12876 TTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 12935
||||||||||||||| ||||| |||||||| |||||||||||| ||||| ||||||
Db 5 TTTTTAGTAGAGATGGGGTTTCACCGTGTTAACCAGGATGGTCTCAATCTCTGGACCTCA 64

Qy 12936 TGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 12981
|||||| || |||||||||||||||||||||||||| |||| ||
Db 65 TGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTATAGGCGTG 110

RESULT 10
AA807640
LOCUS AA807640 103 bp mRNA EST 05-MAR-1998
DEFINITION nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3' similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 774 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
FEATURES Location/Qualifiers
source 1. .103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255473"
/clone_lib="NCI_CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 19 a 27 c 30 g 27 t
ORIGIN

Query Match 0.5%; Score 86.6; DB 38; Length 103;
Best Local Similarity 91.1%; Pred. No. 0.35;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 12881 AGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATC 12940
|||||||||||| ||||| ||||||||||||||||||||||||||||||||||||| ||||||
Db 2 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTTGTGATC 61

Qy 12941 CACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATG 12981
| | |||||||||||||||||||||||||||||| ||
Db 62 CGCTCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGTGTG 102

RESULT 11
B17434
LOCUS B17434 109 bp DNA GSS 04-JUN-1998
DEFINITION 345K2.TVB CIT978SKA1 Homo sapiens genomic clone A-345K02, genomic survey sequence.
ACCESSION B17434
VERSION B17434.1 GI:2125183
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: 345K02.TP 345K02.TPB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="CIT978SKA1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII; CalTech Human BAC Library A1"
BASE COUNT 24 a 30 c 31 g 24 t
ORIGIN

Query Match 0.5%; Score 86.6; DB 84; Length 109;
Best Local Similarity 87.2%; Pred. No. 0.34;
Matches 95; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 8125 TGGCTCACGCCTGTAATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTC 8184
||||||| ||| |||||| |||||||||| ||||||| | |||| |||||| |||||||
Db 1 TGGCTCATACCTATAATCCTAGCACTTTGGGAGGCTGATGTGGGCGGATCACCTGAGGTC 60

Qy 8185 AGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCCACTA 8233
|||||| |||||||||||||||| |||||||||||| ||||| ||||
Db 61 GGGAGTTCGAGACCAGCCTGGCCACCATGGTGAAACCCCGTCTCAACTA 109

```
RESULT  12
B17434/c
LOCUS       B17434         109 bp    DNA             GSS       04-JUN-1998
DEFINITION  345K2.TVB CIT978SKA1 Homo sapiens genomic clone A-345K02, genomic
            survey sequence.
ACCESSION   B17434
VERSION     B17434.1  GI:2125183
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
  TITLE     Use of a BAC End Sequence Database for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: 345K02.TP 345K02.TPB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="A-345K02"
                     /clone_lib="CIT978SKA1"
                     /sex="Female"
                     /cell_type="Fibroblast"
                     /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
                     CalTech Human BAC Library A1"
BASE COUNT       24 a     30 c     31 g     24 t
ORIGIN


  Query Match             0.5%;  Score 86.6;  DB 84;  Length 109;
  Best Local Similarity   87.2%;  Pred. No. 0.34;
  Matches   95;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  11325 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGATGATCTCGAACACCTGACCTCAAGTG 11384
          |||| |||| |||||||||||||| |||||||| || |||||||| || ||||||| |||
Db    109 TAGTTGAGACGGGGTTTCACCATGGTGGCCAGGCTGGTCTCGAACTCCCGACCTCAGGTG 50

Qy  11385 ATCCACCCACCTCAGTCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCA 11433
          |||| ||||| |||| |||||||||||| |||||| |||| |||||||
Db     49 ATCCGCCCACATCAGCCTCCCAAAGTGCTAGGATTATAGGTATGAGCCA 1


RESULT  13
AA244245
LOCUS       AA244245       110 bp    mRNA            EST       20-AUG-1997
DEFINITION  nc07a04.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1007406
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA244245
VERSION     AA244245.1  GI:1875104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 24, 1995 this sequence version replaced gi:634306.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 90.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1007406"
                     /clone_lib="NCI_CGAP_Pr1"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from
                     5,000-10,000 microdissected, histologically normal
                     prostate epithelial cells. Double-stranded cDNA was
                     ligated to EcoRI adaptors, 5 cycles of PCR applied to the
                     cDNA with an adaptor-specific primer, and the resulting
                     PCR product subcloned into pAMP10 by the UDG-cloning
                     method (Life Technologies).  Average insert size is 600
                     bp. NOTE: Not directionally cloned.  This library was
                     constructed by David Krizman."
BASE COUNT       17 a     26 c     28 g     38 t      1 others
ORIGIN


  Query Match             0.5%;  Score 86.6;  DB 30;  Length 110;
  Best Local Similarity   86.4%;  Pred. No. 0.34;
  Matches   95;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy  13412 TTTTTTTTTTGAGATTGAGTCTCATTCTGTCACCCAGGCTGGAGTACAGTGACATGATCT 13471
          |||||||||||||||| ||||||   |||||  |||||||||||||| ||||| ||   |||
Db      1 TTTTTTTTTTGAGATGGAGTCTTGATCTGTTGCCCAGGCTGGAGTGCAGTGGCAGANTCT 60

Qy  13472 TGGCTCATTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGACTCA 13521
          ||||||| |||||||||||||||||||||||||||| |||||| ||| ||||
Db     61 TGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTTCTGCCTCA 110


RESULT  14
AQ003188
LOCUS       AQ003188       110 bp    DNA             GSS       14-APR-1999
DEFINITION  RPCI11-1D10.TPN RPCI-11 Homo sapiens genomic clone RPCI-11-1D10,
            genomic survey sequence.
ACCESSION   AQ003188
VERSION     AQ003188.1  GI:3030392
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
```

```
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: mdadams@tigr.org
           Clones are derived from the human BAC library RPCI-11. For BAC
           library availability, please contact Pieter de Jong
           (pieter@dejong.med.buffalo.edu). Clones may be purchased from
           BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
           Research Genetics (info@resgen.com). BAC end search page:
           http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
           Seq primer: SP6
           Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="GDB:7500081"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-1D10"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       22 a     27 c     26 g     35 t
ORIGIN


  Query Match             0.5%;  Score 86.6;  DB 94;  Length 110;
  Best Local Similarity   87.2%;  Pred. No. 0.34;
  Matches   95;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          ||||||||||||||||||  |  | ||||  || |||| ||||||||||| |||||||
Db      2 TTTTTGTATTTTTAGTAGATACAGGGTTTTACCATGTTGGCCAGGATGGTCCCGATCTCT 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGATTACAG 12976
          ||||||| ||||||||| ||| | ||||||||||||||||||||||||
Db     62 TGACCTCATGATCCACCTGCCGCAGCCTCCCAAAGTGCTGGGATTACAG 110


RESULT  15
AA565533/c
LOCUS       AA565533      107 bp    mRNA            EST       08-SEP-1997
DEFINITION  nk42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA565533
VERSION     AA565533.1  GI:2337172
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393355.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1661   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1016157"
                     /clone_lib="NCI_CGAP_GC2"
                     /tissue_type="germ cell tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Bulk
                     germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCACGAG 3'
                     3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                     Average insert size:  1.2 kb."
BASE COUNT       22 a     34 c     26 g     25 t
ORIGIN


  Query Match             0.5%;  Score 86.2;  DB 35;  Length 107;
  Best Local Similarity   87.9%;  Pred. No. 0.38;
  Matches   94;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy   4311 TGGTGCCCACCTGTAATCCCAGCTACTCAGGAGGCTGAAGCGGGAGGATCACTTGAACCT 4370
           |||||    |||||||||||||||||||||||||||||| ||  ||| ||||||||||||
Db    107 TGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCACGAGAATCACTTGAACCT 48

Qy   4371 GGGAGGCAGAGATTGCAATGAGCTGAGATCATGCCACTGCACTCCAG 4417
           |||||||||| ||||| |||||||||||   ||||||||||||||
Db     47 GGGAGGCAGAGCTTGCAGTGAGCTGAGATTGAGCCACTGCACTCCAG 1


Search completed: June 17, 2000, 02:34:49
Job time: 223056 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:50:11 ; Search time 593.69 Seconds
(without alignments)
3941.217 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTCACTTCTAACCA.....CCATCCAGCTTCTCACTCCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	78.4	0.4	105	4	US-08-481-658B-65	Sequence 65, Appl
	2	78.4	0.4	105	4	US-08-477-504A-65	Sequence 65, Appl
	3	78.4	0.4	105	4	US-08-486-756A-65	Sequence 65, Appl
	4	78.4	0.4	105	4	US-08-485-862B-65	Sequence 65, Appl
	5	78.4	0.4	105	5	US-08-787-739-65	Sequence 65, Appl
c	6	72.8	0.4	105	4	US-08-481-658B-65	Sequence 65, Appl
c	7	72.8	0.4	105	4	US-08-477-504A-65	Sequence 65, Appl
c	8	72.8	0.4	105	4	US-08-486-756A-65	Sequence 65, Appl
c	9	72.8	0.4	105	4	US-08-485-862B-65	Sequence 65, Appl
c	10	72.8	0.4	105	5	US-08-787-739-65	Sequence 65, Appl
	11	63.8	0.4	84	3	US-08-454-557C-91	Sequence 91, Appl
	12	63.8	0.4	84	4	US-08-340-426D-91	Sequence 91, Appl
	13	63.8	0.4	84	4	US-08-450-673C-91	Sequence 91, Appl
	14	63.8	0.4	84	6	PCT-US95-17111A-91	Sequence 91, Appl
	15	60.4	0.3	78	3	US-08-454-557C-70	Sequence 70, Appl
	16	60.4	0.3	78	4	US-08-340-426D-70	Sequence 70, Appl
	17	60.4	0.3	78	4	US-08-450-673C-70	Sequence 70, Appl
	18	60.4	0.3	78	6	PCT-US95-17111A-70	Sequence 70, Appl
	19	59.2	0.3	85	3	US-08-454-557C-92	Sequence 92, Appl
	20	59.2	0.3	85	4	US-08-340-426D-92	Sequence 92, Appl
	21	59.2	0.3	85	4	US-08-450-673C-92	Sequence 92, Appl
	22	59.2	0.3	85	6	PCT-US95-17111A-92	Sequence 92, Appl
c	23	58.2	0.3	84	3	US-08-454-557C-91	Sequence 91, Appl
c	24	58.2	0.3	84	4	US-08-340-426D-91	Sequence 91, Appl
c	25	58.2	0.3	84	4	US-08-450-673C-91	Sequence 91, Appl
c	26	58.2	0.3	84	6	PCT-US95-17111A-91	Sequence 91, Appl
c	27	57.8	0.3	78	3	US-08-454-557C-70	Sequence 70, Appl
c	28	57.8	0.3	78	4	US-08-340-426D-70	Sequence 70, Appl
c	29	57.8	0.3	78	4	US-08-450-673C-70	Sequence 70, Appl
c	30	57.8	0.3	78	6	PCT-US95-17111A-70	Sequence 70, Appl
c	31	57.4	0.3	85	3	US-08-332-766A-44	Sequence 44, Appl
c	32	55.4	0.3	76	3	US-08-454-557C-69	Sequence 69, Appl
c	33	55.4	0.3	76	4	US-08-340-426D-69	Sequence 69, Appl
c	34	55.4	0.3	76	4	US-08-450-673C-69	Sequence 69, Appl
c	35	55.4	0.3	76	6	PCT-US95-17111A-69	Sequence 69, Appl
	36	53.6	0.3	60	3	US-08-454-557C-60	Sequence 60, Appl
	37	53.6	0.3	60	4	US-08-340-426D-60	Sequence 60, Appl
	38	53.6	0.3	60	4	US-08-450-673C-60	Sequence 60, Appl
	39	53.6	0.3	60	6	PCT-US95-17111A-60	Sequence 60, Appl
c	40	52.6	0.3	60	3	US-08-454-557C-57	Sequence 57, Appl
c	41	52.6	0.3	60	4	US-08-340-426D-57	Sequence 57, Appl
c	42	52.6	0.3	60	4	US-08-450-673C-57	Sequence 57, Appl
c	43	52.6	0.3	60	6	PCT-US95-17111A-57	Sequence 57, Appl
c	44	52.2	0.3	98	1	US-08-088-658-42	Sequence 42, Appl
c	45	52.2	0.3	98	4	US-08-471-907A-42	Sequence 42, Appl

ALIGNMENTS

```
RESULT   1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match              0.4%;  Score 78.4;  DB 4;  Length 105;
  Best Local Similarity   84.6%;  Pred. No. 7e-08;
```

```
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          |||||  || ||||||||||||  | ||||| || | || ||||||| || |||| | ||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGAT 12971
          |||||| |||||||||| ||||||||||||||||||||||||||
Db     62 TGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   2
US-08-477-504A-65
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.4%;  Score 78.4;  DB 4;  Length 105;
  Best Local Similarity   84.6%;  Pred. No. 7e-08;
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          |||||  || ||||||||||||  | ||||| || | || ||||||| || |||| | ||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGAT 12971
          |||||| |||||||||| ||||||||||||||||||||||||||
Db     62 TGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   3
US-08-486-756A-65
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.4%;  Score 78.4;  DB 4;  Length 105;
  Best Local Similarity   84.6%;  Pred. No. 7e-08;
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          |||||  || ||||||||||||  | ||||| || | || ||||||| || |||| | ||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGAT 12971
          |||||| |||||||||| ||||||||||||||||||||||||||
Db     62 TGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   4
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
```

```
;       CITY:  Tiburon
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94920
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,862B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;       APPLICATION NUMBER:  US 08/260,190
;       FILING DATE:  15-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.3D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.4%;  Score 78.4;  DB 4;  Length 105;
  Best Local Similarity   84.6%;  Pred. No. 7e-08;
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          |||||  || ||||||||||||  | ||||| || | || |||||| || |||| | ||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGAT 12971
          |||||| |||||||||| ||||||||||||||||||||||||||
Db     62 TGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-787-739-65


  Query Match             0.4%;  Score 78.4;  DB 5;  Length 105;
  Best Local Similarity   84.6%;  Pred. No. 7e-08;
  Matches   88;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12868 TTTTTGTATTTTTAGTAGAGATGGAGTTTCGCCGTGTTAGCCAGGATGGTCTCGATCTCC 12927
          |||||  || ||||||||||||  | ||||| || | || |||||| || |||| | ||||
Db      2 TTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTCC 61

Qy  12928 TGACCTCGTGATCCACCGGCCTCGGCCTCCCAAAGTGCTGGGAT 12971
          |||||| |||||||||| ||||||||||||||||||||||||||
Db     62 TGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   6
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match             0.4%;  Score 72.8;  DB 4;  Length 105;
  Best Local Similarity   86.8%;  Pred. No. 1e-06;
  Matches   92;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy   8140 ATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCA 8199
          ||||||||||||||| |||| ||||| ||  ||||||   |||||||||||||||| ||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48

Qy   8200 GCCTGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||| |||||||||||||||| |||||| ||  |||||
Db     47 GCCTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   7
US-08-477-504A-65/c
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.4%;  Score 72.8;  DB 4;  Length 105;
  Best Local Similarity   86.8%;  Pred. No. 1e-06;
  Matches   92;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy   8140 ATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCA 8199
          ||||||||||||||| |||| ||||| ||  ||||||   |||||||||||||||| ||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48

Qy   8200 GCCTGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||| |||||||||||||||| |||||| ||  |||||
Db     47 GCCTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT   8
US-08-486-756A-65/c
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.4%;  Score 72.8;  DB 4;  Length 105;
  Best Local Similarity   86.8%;  Pred. No. 1e-06;
  Matches   92;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy   8140 ATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCA 8199
          ||||||||||||||| |||| ||||| ||  ||||||   |||||||||||||||| ||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48

Qy   8200 GCCTGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||| |||||||||||||||||| |||||| ||  |||||
Db     47 GCCTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT  9
US-08-485-862B-65/c
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,862B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.4%;  Score 72.8;  DB 4;  Length 105;
  Best Local Similarity   86.8%;  Pred. No. 1e-06;
  Matches   92;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy   8140 ATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCA 8199
          ||||||||||||||| |||| ||||| ||  ||||||   |||||||||||||||| ||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48

Qy   8200 GCCTGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||| |||||||||||||||||| |||||| ||  |||||
Db     47 GCCTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT  10
US-08-787-739-65/c
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  369 Pine Street, Suite 610
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/787,739
;      FILING DATE:  24-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,049
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/486,756
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/477,504
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,658
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,862
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,863
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,077
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-981-2034
;      TELEFAX:  415-981-0332
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-787-739-65
```

```
  Query Match             0.4%;  Score 72.8;  DB 5;  Length 105;
  Best Local Similarity   86.8%;  Pred. No. 1e-06;
  Matches   92;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy   8140 ATCCCAGCACTTTGGCAGGCTGAGGCGGGCAGATCACTTGAGGTCAGGAGTTTGAGACCA 8199
          ||||||||||||||| |||| ||||| ||  ||||||   |||||||||||||||| ||
Db    105 ATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCAC--AAGGTCAGGAGTTTGAGAGCA 48

Qy   8200 GCCTGGCCAACATGGTGAAACCCTGTCTCCACTAAAAATACAAAAA 8245
          |||||||||| ||||||||||||||||||| ||||||  ||  |||||
Db     47 GCCTGGCCAATATGGTGAAACCCTGTCTCTACTAAAGATGTAAAAA 2


RESULT  11
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-91


  Query Match             0.4%;  Score 63.8;  DB 3;  Length 84;
  Best Local Similarity   85.5%;  Pred. No. 6.8e-05;
  Matches   71;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  12899 CCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGGCCTCCC 12958
          || ||||   |||| |||| |||| ||||||||||||||||||| || ||||| |||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  12959 AAAGTGCTGGGATTACAGGCATG 12981
          |||||||||||||||| || ||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT  12
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-91


  Query Match             0.4%;  Score 63.8;  DB 4;  Length 84;
  Best Local Similarity   85.5%;  Pred. No. 6.8e-05;
  Matches   71;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  12899 CCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGGCCTCCC 12958
          || ||||   |||| |||| |||| ||||||||||||||||||| || ||||| |||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  12959 AAAGTGCTGGGATTACAGGCATG 12981
          |||||||||||||||| || ||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT  13
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-91


  Query Match             0.4%;  Score 63.8;  DB 4;  Length 84;
  Best Local Similarity   85.5%;  Pred. No. 6.8e-05;
  Matches   71;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  12899 CCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGGCCTCCC 12958
          || ||||   |||| |||| |||| |||||||||||||||||| || ||||| |||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  12959 AAAGTGCTGGGATTACAGGCATG 12981
          |||||||||||||||| || ||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT  14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-91


  Query Match             0.4%;  Score 63.8;  DB 6;  Length 84;
  Best Local Similarity   85.5%;  Pred. No. 6.8e-05;
  Matches   71;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  12899 CCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCACCGGCCTCGGCCTCCC 12958
          || ||||   |||| |||| |||| |||||||||||||||||| || ||||| |||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  12959 AAAGTGCTGGGATTACAGGCATG 12981
          |||||||||||||||| || ||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT  15
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-70


  Query Match             0.3%;  Score 60.4;  DB 3;  Length 78;
  Best Local Similarity   85.9%;  Pred. No. 0.00034;
  Matches   67;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  13556 ACCATGCCTGGGTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGACCAGG 13615
          || | |||  | |||| ||||||||||||||||||||||||||||| ||||||| | ||||
Db      1 ACAACGCCCAGCTAATATTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTCATCAGG 60

Qy  13616 CTGGTCTCAAACTCCTGA 13633
          ||||| || |||||||||
Db     61 CTGGTGTCGAACTCCTGA 78
```

Search completed: June 17, 2000, 10:50:33
Job time: 251763 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:40:36 ; Search time 13662.6 Seconds
(without alignments)
-1568.842 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
Perfect score: 22034
Sequence: 1 TTCGGAGTGATTTTGGTCAT.....TGTGTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pr1:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vi:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vi:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*
51: gb_pr5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	98.4	0.4	108	10	HSLDLRN2	X05250 Human LDL-r
c	2	90.8	0.4	107	9	HUMALCE162	M87924 Human carci
	3	90.4	0.4	108	10	HSLDLRN2	X05250 Human LDL-r
c	4	89.6	0.4	108	11	HSU67803	U67803 Human small
	5	87.8	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
c	6	87.8	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
c	7	87.2	0.4	103	9	HUMALCE221	M87896 Human carci
c	8	82.4	0.4	104	9	HUMALCE272	M87899 Human carci
c	9	81.4	0.4	108	10	HSLDLRD1	X05249 Human LDL-r
	10	81.4	0.4	108	10	HSLDLRD2	X05251 Human LDL-r
c	11	81.6	0.4	108	11	HSU67804	U67804 Human small
	12	79.6	0.4	107	9	HUMALCE162	M87924 Human carci
	13	78.8	0.4	103	9	HUMALCE221	M87896 Human carci
c	14	78.8	0.4	110	9	HUMALCE43	M87900 Human carci
	15	76.8	0.3	108	10	HSLDLI12	X05248 Human LDL-r
c	16	76.2	0.3	101	10	S79560	S79560 HRX (intron
c	17	75.8	0.3	108	9	HUMD1D03M5	D16965 Human HepG2
	18	75	0.3	103	13	HS8IC8R	X57789 Human seque
c	19	75.2	0.3	108	11	HSU67808	U67808 Human small
c	20	74.4	0.3	110	11	HSU67807	U67807 Human small
	21	73.6	0.3	106	13	G32743	G32743 A009P31 Hum
	22	72.6	0.3	108	9	HUMD1D03M5	D16965 Human HepG2
	23	72	0.3	90	9	HUMLDLRFL	K03555 Human low d
c	24	72	0.3	91	13	HUMUT8164A	L30244 Human STS U
c	25	70.4	0.3	107	11	HSU67806	U67806 Human small
c	26	70.4	0.3	108	10	HSLDLI12	X05248 Human LDL-r
	27	69.4	0.3	100	13	HUMUT931A	L31299 Human STS U
	28	69.6	0.3	104	9	HUMALCE272	M87899 Human carci
c	29	68.4	0.3	79	10	S73203	S73203 ALL-1 (tand
	30	68.4	0.3	95	13	HUMUT8002B	L30176 Human STS U
	31	68.2	0.3	102	13	G32906	G32906 A009W09 Hum
	32	67	0.3	84	5	AR051521	AR051521 Sequence
	33	67	0.3	108	13	G43535	G43535 WIAF-2393-S
	34	66	0.3	91	13	HUMUT8164A	L30244 Human STS U
	35	66.2	0.3	100	9	HUMGALNSA	D45223 Human GALNS
	36	65.6	0.3	80	9	HUMBRKFAE	M36135 Human alpha
	37	65.8	0.3	97	9	HUMLDLRA2	M14180 Human low d
	38	65.2	0.3	95	10	HSSTHPKIB	X66361 H.sapiens m
c	39	64.4	0.3	100	13	HUMUT931A	L31299 Human STS U
c	40	64	0.3	80	9	HUMBRKFAE	M36135 Human alpha
	41	64	0.3	107	13	G32919	G32919 A009W27 Hum
c	42	63.4	0.3	90	10	HSU19407	U19407 Human isola
c	43	63.6	0.3	95	13	HUMUT8002B	L30176 Human STS U
c	44	63.8	0.3	106	13	G32743	G32743 A009P31 Hum
c	45	63.8	0.3	108	13	G32614	G32614 A009K21 Hum

ALIGNMENTS

```
RESULT   1
HSLDLRN2/c
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.4%;  Score 98.4;  DB 10;  Length 108;
  Best Local Similarity   94.4%;  Pred. No. 3.2e-06;
  Matches  102;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy  14586 CTTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTTCCA 14645
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    108 CTCGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCG 49

Qy  14646 AGTAGCTGGGATTACAGGCACCCACTACCACGCCTGGCTAATTTTTGT 14693
          |||||||||||||||||||||  | |||||||||||||||||||||
Db     48 AGTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   2
HUMALCE162/c
LOCUS       HUMALCE162    107 bp ss-RNA           PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION   M87924
VERSION     M87924.1  GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       28 a     30 c     35 g     14 t
ORIGIN


  Query Match             0.4%;  Score 90.8;  DB 9;  Length 107;
  Best Local Similarity   93.1%;  Pred. No. 4.9e-05;
  Matches   95;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy  11779 TTTTTGAGACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCT 11838
          |||||||||||||||||| ||||||| ||||||||||||||||||||||||||||||||
Db    106 TTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCT 47

Qy  11839 CACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCC 11880
          |||||||| |||||||||||||||| |  ||||| |||||
Db     46 CACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTTCTGCC 5


RESULT   3
HSLDLRN2
LOCUS       HSLDLRN2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 14 fragment (normal gene).
ACCESSION   X05250
VERSION     X05250.1  GI:34337
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     See X05252 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     23 c     39 g     18 t
ORIGIN


  Query Match             0.4%;  Score 90.4;  DB 10;  Length 108;
  Best Local Similarity   89.8%;  Pred. No. 5.6e-05;
  Matches   97;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   4976 ACAAAAATTAGCCAGGTGTGGTGGCAGGCACCTGTAATCCCAGCTATTCAGGAGGCTGAG 5035
           |||||||||||||||  ||||||||||  ||||||||||||||| || ||||||||||
Db      1 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 60

Qy   5036 ACAGGAGAATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAG 5083
           |||||||| ||||||||||||||||  ||||||||| |||||| ||
Db     61 GCAGGAGAATTGCTTGAACCCAGGAGGCAGAGGTTGCAGTGAGCCGAG 108


RESULT   4
HSU67803/c
LOCUS       HSU67803      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67803
VERSION     U67803.1  GI:2289917
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
```

```
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu2"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       23 a     39 c     30 g     16 t
ORIGIN


  Query Match             0.4%;  Score 89.6;  DB 11;  Length 108;
  Best Local Similarity   95.8%;  Pred. No. 7.5e-05;
  Matches   92;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy  10437 GTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 10496
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     97 GTAGAGACGGGGTTTCACCTTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 38

Qy  10497 GCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGG 10532
          |||| |||  |||||||||||||||||||||||||||
Db     37 GCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 2


RESULT   5
HSLDLRD1
LOCUS       HSLDLRD1      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION   X05249
VERSION     X05249.1  GI:34335
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05248 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     misc_feature    1. .108
                     /note="deletion junction region intron 12/ intron 15"
BASE COUNT       20 a     40 c     20 g     28 t
ORIGIN


  Query Match             0.4%;  Score 87.8;  DB 10;  Length 108;
  Best Local Similarity   88.8%;  Pred. No. 0.00014;
  Matches   95;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  14587 TTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTTCCAA 14646
          | | |||||  |||||||||||||||||||||||| | ||| |||||||||||||| || |
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 61

Qy  14647 GTAGCTGGGATTACAGGCACCCACTACCACGCCTGGCTAATTTTTGT 14693
          |||||||||||||||||||  | |||||||||||||||||||||||
Db     62 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 108


RESULT   6
HSLDLRD2/c
LOCUS       HSLDLRD2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION   X05251
VERSION     X05251.1  GI:34336
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     20 c     40 g     20 t
ORIGIN


  Query Match             0.4%;  Score 87.8;  DB 10;  Length 108;
  Best Local Similarity   88.8%;  Pred. No. 0.00014;
  Matches   95;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  14587 TTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCTTCCAA 14646
          | | |||||  |||||||||||||||||||||||| | ||| |||||||||||||| || |
Db    107 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCCGA 48

Qy  14647 GTAGCTGGGATTACAGGCACCCACTACCACGCCTGGCTAATTTTTGT 14693
          |||||||||||||||||||  | |||||||||||||||||||||||
Db     47 GTAGCTGGGATTACAGGCACCTGCCACCACGCCTGGCTAATTTTTGT 1


RESULT   7
HUMALCE221/c
LOCUS       HUMALCE221    103 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION   M87896
VERSION     M87896.1  GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
```

```
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       25 a     27 c     33 g     18 t
ORIGIN


  Query Match             0.4%;  Score 87.2;  DB 9;  Length 103;
  Best Local Similarity   92.0%;  Pred. No. 0.00018;
  Matches   92;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy  11812 CTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGAT 11871
          |||||||||| |||| ||||||||||||||||||||||||||||||||||||||||| |||
Db    103 CTGGAGTGCAATGGCACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGAT 44

Qy  11872 TCTCCTGCCTCAGACTCCCGAGTAGCTGGGATTACAGTCA 11911
          |||||||||| ||  ||||| ||||||||||||||| ||
Db     43 TCTCCTGCCTTAGCTTCCCGTGTAGCTGGGATTACAGGCA 4


RESULT   8
HUMALCE272/c
LOCUS       HUMALCE272    104 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION   M87899
VERSION     M87899.1  GI:174875
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 104)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .104
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       22 a     26 c     37 g     19 t
ORIGIN


  Query Match             0.4%;  Score 82.4;  DB 9;  Length 104;
  Best Local Similarity   89.0%;  Pred. No. 0.001;
  Matches   89;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  10454 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAGCCTCCC 10513
          |||||||||||||| |||||| || |||||| |||||  |||  ||| ||| ||||||||
Db    104 CCATGTTAGCCAGGCTGGTCTTGAACTCCTGGCCTCGCAATCTTCCCTCCTCAGCCTCCC 45

Qy  10514 AAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCC 10553
          |||||||||||||||||| ||||||||||||||||||||
Db     44 AAAGTGCTGGGATTACAGGAGTGAGCCACCGCGCCCGGCC 5


RESULT   9
HSLDLRD1/c
LOCUS       HSLDLRD1      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION   X05249
VERSION     X05249.1  GI:34335
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05248 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     misc_feature    1. .108
                     /note="deletion junction region intron 12/ intron 15"
BASE COUNT       20 a     40 c     20 g     28 t
ORIGIN


  Query Match             0.4%;  Score 81.4;  DB 10;  Length 108;
  Best Local Similarity   85.0%;  Pred. No. 0.0014;
  Matches   91;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   4976 ACAAAAATTAGCCAGGTGTGGTGGCAGGCACCTGTAATCCCAGCTATTCAGGAGGCTGAG 5035
           |||||||||||||||| |||||||||||  ||||||||||||||| || ||||||||||
Db    108 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 49

Qy   5036 ACAGGAGAATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAA 5082
           ||||| ||| | |||||||||||||  |||||||   |||| | |
Db     48 GCAGGAAAATGGTTTGAACCCAGGAGGCAGAGGTTGTGGTGAGGCGA 2


RESULT  10
HSLDLRD2
LOCUS       HSLDLRD2      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION   X05251
VERSION     X05251.1  GI:34336
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     *source: hypercholesterol aemia
            See X05250 for corresponding normal gene sequence
            In the defective LDL-receptor gene the deletion ocurred between two
            alu-repetitive sequences, that are in the same direction, the
            deletion eliminates exons 13 and 14 and changes the reading frame
            of the resulting spliced mRNA.
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
```

```
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="blood leukocytes from a patient with familial"
     intron          1. .108
                     /note="intron XIV fragment"
BASE COUNT       28 a     20 c     40 g     20 t
ORIGIN


  Query Match             0.4%;  Score 81.4;  DB 10;  Length 108;
  Best Local Similarity   85.0%;  Pred. No. 0.0014;
  Matches   91;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy    4976 ACAAAAATTAGCCAGGTGTGGTGGCAGGCACCTGTAATCCCAGCTATTCAGGAGGCTGAG 5035
           |||||||||||||||| |||||||||||  |||||||||||||||| || ||||||||||
Db       1 ACAAAAATTAGCCAGGCGTGGTGGCAGGTGCCTGTAATCCCAGCTACTCGGGAGGCTGAG 60

Qy    5036 ACAGGAGAATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAA 5082
           ||||| ||| | |||||||||||||  |||||||   |||| | |
Db      61 GCAGGAAAATGGTTTGAACCCAGGAGGCAGAGGTTGTGGTGAGGCGA 107


RESULT  11
HSU67804/c
LOCUS       HSU67804      108 bp    RNA             PRI       01-AUG-1997
DEFINITION  Human small cytoplasmic Alu transcript.
ACCESSION   U67804
VERSION     U67804.1  GI:2289918
KEYWORDS    Alu.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     cDNAs derived from primary and small cytoplasmic Alu (scAlu)
            transcripts
  JOURNAL   J. Mol. Biol. 271 (2), 222-234 (1997)
  MEDLINE   97415756
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TscAlu3"
     repeat_region   1. .108
                     /note="scAlu"
                     /rpt_family="Alu"
                     /rpt_type=dispersed
BASE COUNT       26 a     38 c     26 g     18 t
ORIGIN


  Query Match             0.4%;  Score 81.6;  DB 11;  Length 108;
  Best Local Similarity   90.6%;  Pred. No. 0.0013;
  Matches   87;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   10437 GTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCC 10496
           | | |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      97 GGAAAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTTGTGATCC 38

Qy   10497 GCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGG 10532
           ||| | |  |||| ||||||||||||||||||||
Db      37 TCCCGCTTTGGCCTTCCAAAGTGCTGGGATTACAGG 2


RESULT  12
HUMALCE162
LOCUS       HUMALCE162    107 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION   M87924
VERSION     M87924.1  GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       28 a     30 c     35 g     14 t
ORIGIN


  Query Match             0.4%;  Score 79.6;  DB 9;  Length 107;
  Best Local Similarity   86.3%;  Pred. No. 0.0027;
  Matches   88;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    5035 GACAGGAGAATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACT 5094
           | ||| ||||| || ||||||| ||||| |||| ||||| |||||| |||||  ||||||
Db       5 GGCAGAAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACT 64

Qy    5095 GCACTCCAGCCTGGGCGACAGAGTAAGACTCCGTTTCAAAAA 5136
           |||||||||||||||||||||||  |||||||| |||||||
Db      65 GCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCAAAAA 106


RESULT  13
HUMALCE221
LOCUS       HUMALCE221    103 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION   M87896
VERSION     M87896.1  GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       25 a     27 c     33 g     18 t
ORIGIN


  Query Match             0.4%;  Score 78.8;  DB 9;  Length 103;
  Best Local Similarity   87.8%;  Pred. No. 0.0036;
  Matches   86;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    5006 CCTGTAATCCCAGCTATTCAGGAGGCTGAGACAGGAGAATCGCTTGAACCCAGGAGGTGG 5065
```

```
          |||||||||||||||  | ||| ||| || ||||||||||||||||||| ||||| ||
Db      6 CCTGTAATCCCAGCTACACGGGAAGCTAAGGCAGGAGAATCGCTTGAACCCGGGAGGCGG 65

Qy   5066 AGGTTGCATTGAGCCAAGATCATGCCACTGCACTCCAG 5103
          |||||||| |||||| ||||| ||||| ||||||||||
Db     66 AGGTTGCAGTGAGCCGAGATCGTGCCATTGCACTCCAG 103


RESULT  14
HUMALCE43/c
LOCUS       HUMALCE43     110 bp ss-RNA             PRI       15-APR-1994
DEFINITION  Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION   M87900
VERSION     M87900.1  GI:174876
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
  TITLE     Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
  JOURNAL   J. Mol. Biol. (1992) In press
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="NTera2D1"
                     /dev_stage="embryo"
                     /sex="male"
                     /tissue_type="carcinoma"
BASE COUNT       27 a     31 c     34 g     18 t
ORIGIN


  Query Match             0.4%;  Score 78.8;  DB 9;  Length 110;
  Best Local Similarity   84.0%;  Pred. No. 0.0036;
  Matches   89;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  10448 GTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAG 10507
          |||||  |||||||||||||| |||||| || || |||| ||||  |||| ||  |||  |
Db    110 GTTTCGTCATGTTAGCCAGGCTGGTCTTGAACTACTGAGCTCGCAATCCTCCTGCCTTGG 51

Qy  10508 CCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCC 10553
          |||||||||||||| ||||||  ||||||||||||||||||||||
Db     50 CCTCCCAAAGTGCCGGGATTGTAGGTGTGAGCCACCGCGCCCGGCC 5


RESULT  15
HSLDLI12
LOCUS       HSLDLI12      108 bp    DNA             PRI       20-MAY-1992
DEFINITION  Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low
            density lipoprotein.
ACCESSION   X05248
VERSION     X05248.1  GI:34334
KEYWORDS    Alu repetitive sequence; low density lipoprotein receptor;
            repetitive sequence.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
            Williamson,R. and Humphries,S.
  TITLE     Unequal crossing-over between two alu-repetitive DNA sequences in
            the low-density-lipoprotein-receptor gene. A possible mechanism for
            the defect in a patient with familial hypercholesterolaemia
  JOURNAL   Eur. J. Biochem. 164 (1), 77-81 (1987)
  MEDLINE   87161901
COMMENT     see X05249 for deletion junction
            Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    complement(<1. .65)
                     /note="Alu repeat"
     intron          1. .108
                     /note="intron XII fragment"
BASE COUNT       21 a     38 c     20 g     29 t
ORIGIN


  Query Match             0.3%;  Score 76.8;  DB 10;  Length 108;
  Best Local Similarity   83.7%;  Pred. No. 0.0074;
  Matches   87;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy  18268 TCGGCTCACAGCAACCTCCACCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGA 18327
          ||| |||||  |||||||  ||||||||||||||   ||| |||||||||||||||||| |
Db      2 TCGCCTCACCACAACCTCTGCCTCCTGGGTTCAAACCATTTTCCTGCCTCAGCCTCCTTA 61

Qy  18328 GTAGCTGGGATTACAGACGTGTGTCACCACACCTGGCTAATTTT 18371
          ||||||||||||||  | |||| |||||| || |||| |||||
Db     62 GTAGCTGGGATTACAAGCATGTGCCACCACGCCCGGCTGATTTT 105


Search completed: June 17, 2000, 17:34:18
Job time: 276942 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 11:43:23 ; Search time 455.92 Seconds
(without alignments)
12091.447 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
Perfect score: 22034
Sequence: 1 TTCGGAGTGATTTTGGTCAT.....TGTGTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	67.6	0.3	108	1	X12095	Human biallelic po
c	2	67.2	0.3	100	1	T24892	Human gene signatu
	3	66	0.3	108	1	T26828	Human gene signatu
	4	63.4	0.3	108	1	T25009	Human gene signatu
c	5	62.6	0.3	91	1	T25854	Human gene signatu
c	6	61.8	0.3	108	1	T25009	Human gene signatu
	7	60.8	0.3	100	1	T24892	Human gene signatu
	8	60	0.3	103	1	T26213	Human gene signatu
	9	59.6	0.3	97	1	T26728	Human gene signatu
c	10	59.4	0.3	103	1	T26213	Human gene signatu
	11	58	0.3	103	1	T20927	Human gene signatu
c	12	56.4	0.3	107	1	T20373	Human gene signatu
c	13	56.6	0.3	108	1	T26828	Human gene signatu
c	14	54.8	0.2	87	1	T21566	Human gene signatu
	15	54.8	0.2	93	1	T22572	Human gene signatu
	16	55	0.2	95	1	T23131	Human gene signatu
c	17	54.4	0.2	93	1	T25688	Human gene signatu
	18	54.2	0.2	93	1	T24259	Human gene signatu
c	19	54.2	0.2	93	1	T24259	Human gene signatu
c	20	54.2	0.2	95	1	T23131	Human gene signatu
c	21	54.4	0.2	106	1	Q95210	Simple tandem repe
	22	54	0.2	100	1	X12087	Human biallelic po
	23	54	0.2	100	1	X12085	Human biallelic po
c	24	53.6	0.2	69	1	Q29016	Probe to internal
	25	53.6	0.2	110	1	T26288	Human gene signatu
	26	53	0.2	91	1	T25854	Human gene signatu
	27	52.8	0.2	100	1	X12086	Human biallelic po
	28	53	0.2	109	1	T23895	Human gene signatu
	29	50.8	0.2	70	1	N60231	Normal chromosome
c	30	50	0.2	69	1	T24175	Human gene signatu
c	31	50	0.2	81	1	T24093	Human gene signatu
c	32	50	0.2	93	1	T22572	Human gene signatu
	33	50	0.2	105	1	T23942	Human gene signatu
c	34	49	0.2	65	1	T24893	Human gene signatu
c	35	49	0.2	91	1	T26410	Human gene signatu
	36	48.4	0.2	75	1	T22841	Human gene signatu
	37	48.6	0.2	99	1	T20931	Human gene signatu
c	38	48.4	0.2	99	1	T23728	Human gene signatu
c	39	48.2	0.2	101	1	T24667	Human gene signatu
c	40	47.4	0.2	65	1	T25588	Human gene signatu
	41	47.6	0.2	94	1	T26403	Human gene signatu
c	42	47.6	0.2	94	1	T26403	Human gene signatu
c	43	47.8	0.2	99	1	T20931	Human gene signatu
c	44	47.8	0.2	103	1	T20927	Human gene signatu
	45	47	0.2	87	1	V05740	Nucleotide sequenc

ALIGNMENTS

RESULT 1
X12095
ID X12095 standard; DNA; 108 BP.
AC X12095;
DT 30-MAR-1999 (first entry)
DE Human biallelic polymorphic DNA fragment TIGR-A003M18a.
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW treatment; marker; ss.
OS Homo sapiens.
PN WO9820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1; Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.3%; Score 67.6; DB 1; Length 108;
Best Local Similarity 82.6%; Pred. No. 0.065;
Matches 76; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

```
Qy  14691 TGTATTTCTGGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGAC 14750
          ||| ||| | ||||||| | ||||| | :|||||||||| ||||||||||||||||||
Db      1 TGTCTTTTTTGTAGAGATGAGGTTTTCCTRTGTTGGCCAGGATGGTCTCGAACTCCTGAC 60

Qy  14751 CTCAAGTGATCCACAATCCTTGGCCTCCCAAA 14782
            ||||||||||      ||||||||||||||
Db     61 TTCAAGTGATCCGTCTGCCTTGGCCTCCCAAA 92
```

```
RESULT   2
T24892/c
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;


  Query Match             0.3%;  Score 67.2;  DB 1;  Length 100;
  Best Local Similarity   78.8%;  Pred. No. 0.074;
  Matches   78;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy  11773 TATTTATTTTTGAGACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATC 11832
          | | | | ||| | || |||| ||||||||||||||||| ||||||||  || || |||
Db    100 TTTGTTTGTTTCAAACAGAGTGTCACTCTGTCACCCAGGCNGGAGTGCAANGGTGCAATC 41

Qy  11833 TCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGAT 11871
          || ||| | |||||  || ||||||| |||||||| |||
Db     40 TCAGCTNATTGCAAATTCTGCCTCCCAGGTTCAAGCGAT 2


RESULT   3
T26828
ID   T26828 standard; cDNA to mRNA; 108 BP.
AC   T26828;
DT   14-NOV-1996  (first entry)
DE   Human gene signature HUMGS09078.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2182; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    18 A;    33 C;    23 G;    28 T;


  Query Match             0.3%;  Score 66;  DB 1;  Length 108;
  Best Local Similarity   90.8%;  Pred. No. 0.11;
  Matches   69;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1728 ATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 1787
           |||||||||||||| |||||||||  |||||||||| ||||||||| |||||||| ||
Db      2 ATCTCCTGACCTCGTCATCCGCCCGTNTCGGCCTCCCATAGTGCTGGGNTTACAGGCATG 61

Qy   1788 AGCCACCGCGCCCGGC 1803
           ||||||| ||||||||
Db     62 AGCCACCACGCCCGGC 77


RESULT   4
T25009
ID   T25009 standard; cDNA to mRNA; 108 BP.
AC   T25009;
DT   07-NOV-1996  (first entry)
DE   Human gene signature HUMGS07131.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1748; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
```

```
SQ   Sequence   108 BP;    34 A;    31 C;    26 G;    15 T;


  Query Match             0.3%;  Score 63.4;  DB 1;  Length 108;
  Best Local Similarity   73.8%;  Pred. No. 0.26;
  Matches   79;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy   5044 ATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACTGCACTCCAG 5103
           ||||| ||| |||| ||||   ||| |||| ||||||| | ||| ||||||| | |||||
Db      2 ATCGCCTGAGCCCATGAGGCCAAGGCTGCAGTGAGCCATGGTCACGCCACTGNATTCCAG 61

Qy   5104 CCTGGGCGACAGAGTAAGACTCCGTTTCAAAAACAAAAAACCCAAGA 5150
           |||| | ||||||| ||||| | |||  ||| |  || |||   | |
Db     62 CCTGAGTGACAGAGCAAGACCCTGTTGAAAACAACAACAACANCAAA 108


RESULT   5
T25854/c
ID   T25854 standard; cDNA to mRNA; 91 BP.
AC   T25854;
DT   22-OCT-1996  (first entry)
DE   Human gene signature HUMGS08084.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1944; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   91 BP;    18 A;    22 C;    28 G;    18 T;


  Query Match             0.3%;  Score 62.6;  DB 1;  Length 91;
  Best Local Similarity   79.8%;  Pred. No. 0.34;
  Matches   71;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  11783 TGAGACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACT 11842
          |||||| |  |||||| |||||||| ||||||||| ||||  | || ||||| |||||||
Db     90 TGAGACAGNNTCTCACGCTGTCACCNAGGCTGGAGCGCAGGAGTGCCATCTCAGCTCACT 31

Qy  11843 GCAACCTCCGCCTCCCGGGTTCAAGTGAT 11871
            |||| | ||||||  || |||||||||
Db     30 TGAACCNCTGCCTCCTAGGCTCAAGTGAT 2


RESULT   6
T25009/c
ID   T25009 standard; cDNA to mRNA; 108 BP.
AC   T25009;
DT   07-NOV-1996  (first entry)
DE   Human gene signature HUMGS07131.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1748; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    34 A;    31 C;    26 G;    15 T;


  Query Match             0.3%;  Score 61.8;  DB 1;  Length 108;
  Best Local Similarity   72.9%;  Pred. No. 0.44;
  Matches   78;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy  10253 TTTTTTTTTTTTTTTTTTTTAGACAGAGTCTTACTCTGTCGCCCGGGCTGGAGTGCAGTG 10312
          |||  | || || || ||||  |||| ||||| ||||||| | | ||||||| | |||||
Db    108 TTTGNTGTTGTTGTTGTTTTCAACAGGGTCTTGCTCTGTCACTCAGGCTGGAATNCAGTG 49

Qy  10313 GCTCGATCTTAGCTCACTGAAAGCTTTGCCTCCCGGGTTCATGCCAT 10359
          ||  || | | |||||||| |  ||| |||||  ||| ||| || ||
Db     48 GCGTGACCATGGCTCACTGCAGCCTTGGCCTCATGGGCTCAGGCGAT 2


RESULT   7
T24892
ID   T24892 standard; cDNA to mRNA; 100 BP.
AC   T24892;
DT   05-NOV-1996  (first entry)
DE   Human gene signature HUMGS06998.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
```

```
PS   Claim 1; Page 1720; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   100 BP;    28 A;    22 C;    25 G;    22 T;


  Query Match             0.3%;  Score 60.8;  DB 1;  Length 100;
  Best Local Similarity   74.7%;  Pred. No. 0.61;
  Matches   74;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy   5044 ATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACTGCACTCCAG 5103
            |||||||||||  |||||  ||  ||||| | |||  ||||    ||  |||||||| |
Db      2 ATCGCTTGAACCTGGGAGGCAGAATTTGCAATNAGCTGAGATTGCACCNTTGCACTCCNG 61

Qy   5104 CCTGGGCGACAGAGTAAGACTCCGTTTCAAAAACAAAAA 5142
          |||||| |||||||| | |||| |||| ||| | | |||
Db     62 CCTGGGTGACAGAGTGACACTCTGTTTGAAACAAACAAA 100


RESULT   8
T26213
ID   T26213 standard; cDNA to mRNA; 103 BP.
AC   T26213;
DT   13-NOV-1996  (first entry)
DE   Human gene signature HUMGS08452.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2029; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    33 A;    21 C;    25 G;    23 T;


  Query Match             0.3%;  Score 60;  DB 1;  Length 103;
  Best Local Similarity   75.0%;  Pred. No. 0.79;
  Matches   75;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy   5044 ATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACTGCACTCCAG 5103
          ||| |||||  ||||||| | |  ||| || ||||| | |||    ||||||||||||||
Db      2 ATCACTTGAGTCCAGGAGTTTGGTGTTACAGTGAGCTATGATGGCACCACTGCACTCCAG 61

Qy   5104 CCTGGGCGACAGAGTAAGACTCCGTTTCAAAAACAAAAAA 5143
          ||||||| ||||||||||||    || |  || | ||||||
Db     62 CCTGGGCCACAGAGTAAGAACATGTCTTTAAGAAAAAAAA 101


RESULT   9
T26728
ID   T26728 standard; cDNA to mRNA; 97 BP.
AC   T26728;
DT   22-OCT-1996  (first entry)
DE   Human gene signature HUMGS08978.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2158; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   97 BP;    19 A;    27 C;    20 G;    28 T;


  Query Match             0.3%;  Score 59.6;  DB 1;  Length 97;
  Best Local Similarity   81.0%;  Pred. No. 0.91;
  Matches   68;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy  12004 GATCTACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCAG 12063
          ||||| ||||||| ||||||||| |||||||||||||||||||||||||||||||||| |
Db      1 GATCTGCCCACCTNGGCCTCCCAGAGTGCTGGGATTACAGGCATGAGCCACTGCGCCCGG 60

Qy  12064 CCCTGCTTGTCTTTTTATTTTATA 12087
            |     |   | ||| |||| ||
Db     61 NCTGTACTAAGTCTTTTTTTTTA 84


RESULT  10
T26213/c
ID   T26213 standard; cDNA to mRNA; 103 BP.
AC   T26213;
DT   13-NOV-1996  (first entry)
DE   Human gene signature HUMGS08452.
```

```
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2029; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    33 A;    21 C;    25 G;    23 T;

  Query Match             0.3%;  Score 59.4;  DB 1;  Length 103;
  Best Local Similarity   74.3%;  Pred. No. 0.97;
  Matches   75;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy  10259 TTTTTTTTTTTTTTAGACAGAGTCTTACTCTGTCGCCCGGGCTGGAGTGCAGTGGCTCGA 10318
          ||||||||| ||  |||||   |||||||||| |||| ||||||||||||||  | |
Db    102 TTTTTTTTTCTTAAAGACATGTTCTTACTCTGTGGCCCAGGCTGGAGTGCAGTGGTGCCA 43

Qy  10319 TCTTAGCTCACTGAAAGCTTTGCCTCCCGGGTTCATGCCAT 10359
          || |||||||||| ||       |||| ||  ||| |  ||
Db     42 TCATAGCTCACTGTAACACCAAACTCCTGGACTCAAGTGAT 2


RESULT  11
T20927
ID   T20927 standard; cDNA to mRNA; 103 BP.
AC   T20927;
DT   24-JUL-1996  (first entry)
DE   Human gene signature HUMGS02180.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 758-759; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   103 BP;    22 A;    27 C;    21 G;    31 T;

  Query Match             0.3%;  Score 58;  DB 1;  Length 103;
  Best Local Similarity   73.0%;  Pred. No. 1.5;
  Matches   73;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy  18304 GATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGACGTGTGTCACCACACCTGG 18363
          ||| ||||  | ||  |||||  |||||||| |  |||||  ||||| |||||   |  |
Db      1 GATCCTCCCACTTCCACCTCCCAAGTAGCTGTGGCTACAGGTGTGTGCCACCATGTCCAG 60

Qy  18364 CTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCGTG 18403
          || ||||| ||||||| ||||| |||||  |||| || ||
Db     61 CTGATTTTNGTATTTTNAGTAGGGACAGTATTTCTCCATG 100


RESULT  12
T20373/c
ID   T20373 standard; cDNA to mRNA; 107 BP.
AC   T20373;
DT   19-JUL-1996  (first entry)
DE   Human gene signature HUMGS01525.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 623; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   107 BP;    26 A;    29 C;    17 G;    29 T;

  Query Match             0.3%;  Score 56.4;  DB 1;  Length 107;
  Best Local Similarity   72.6%;  Pred. No. 2.6;
  Matches   69;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;
```

```
Qy   4817 ATTCCAAAGCTTCTTTTAGAAATAATATTGCTGGGCCAGGCATGGTGGCTCATGCCTGTA 4876
          | |  | ||| |   |||  || || |||    ||||||| ||||| | ||||||||||
Db     95 AGTTNAGAGCCTTAATTANGAAAAACATTAGAAGGCCAGGNATGGTAGNTCATGCCTGTA 36

Qy   4877 ATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGATC 4911
          ||||||| |||||||||||  |||   || ||||
Db     35 ATTCCAGNACTTTGGGAGGCTTAGGTGCGCGGATC 1


RESULT  13
T26828/c
ID   T26828 standard; cDNA to mRNA; 108 BP.
AC   T26828;
DT   14-NOV-1996  (first entry)
DE   Human gene signature HUMGS09078.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 2182; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   108 BP;    18 A;    33 C;    23 G;    28 T;


  Query Match             0.3%;  Score 56.6;  DB 1;  Length 108;
  Best Local Similarity   84.9%;  Pred. No. 2.4;
  Matches   62;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   4851 GCCAGGCATGGTGGCTCATGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGAT 4910
          ||| ||| ||||||||||||||||||||  |||||||| ||||||||||  | ||| |||
Db     77 GCCGGGCGTGGTGGCTCATGCCTGTAANCCCAGCACTATGGGAGGCCGANACGGGCGGAT 18

Qy   4911 CATGAGGTCAGGA 4923
            | ||||||||||
Db     17 GACGAGGTCAGGA 5


RESULT  14
T21566/c
ID   T21566 standard; cDNA to mRNA; 87 BP.
AC   T21566;
DT   03-AUG-1996  (first entry)
DE   Human gene signature HUMGS02944.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 914; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   87 BP;    35 A;    21 C;    16 G;    13 T;


  Query Match             0.2%;  Score 54.8;  DB 1;  Length 87;
  Best Local Similarity   77.4%;  Pred. No. 4.4;
  Matches   65;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy  14678 CCTGGCTAATTTTTGTATTTCTGGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTC 14737
          | ||||| |||| ||||||| | ||| ||| |||||||| |||||||   | |||||||
Db     84 CTTGGCTNATTTCTGTATTTTTTGTAAAGATGGGGTTTCGCCATGTTCCTCGGGCTGGTT 25

Qy  14738 TCGAACTCCTGACCTCAAGTGATC 14761
          |  ||||||||   ||||| ||||
Db     24 TTAAACTCCTGGGNTCAAGCGATC 1


RESULT  15
T22572
ID   T22572 standard; cDNA to mRNA; 93 BP.
AC   T22572;
DT   01-OCT-1996  (first entry)
DE   Human gene signature HUMGS04188.
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
OS   Homo sapiens.
PN   WO9514772-A1.
PD   01-JUN-1995.
PF   11-NOV-1994; J01916.
PR   12-NOV-1993; JP-355504.
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
PI   Matsubara K,  Okubo K;
DR   WPI; 95-206931/27.
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
PS   Claim 1; Page 1159; 2245pp; Japanese.
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
```

```
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
SQ   Sequence   93 BP;    22 A;    23 C;    24 G;    22 T;


  Query Match             0.2%;  Score 54.8;  DB 1;  Length 93;
  Best Local Similarity   73.9%;  Pred. No. 4.4;
  Matches   68;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy   5044 ATCGCTTGAACCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACTGCACTCCAG 5103
          ||| ||||| || ||||||| |||| |||| ||||| | | |||  ||||||||||||||
Db      2 ATCCCTTGAGCCGAGGAGGTTGAGGCTGCAGTGAGCTATGGTCACACCACTGCACTCCAG 61

Qy   5104 CCTGGGCGACAGAGTAAGACTCCGTTTCAAAA 5135
          | | ||  |    | ||||| | ||||  |||
Db     62 CNTTGGTTATNATGCAAGACCCTGTTTACAAA 93


Search completed: June 17, 2000, 18:05:14
Job time: 277625 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 02:34:49 ; Search time 6489.89 Seconds
(without alignments)
13761.207 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
Perfect score: 22034
Sequence: 1 TTCGGAGTGATTTTGGTCAT.....TGTGTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gss1:*
83: gb_gss2:*
84: gb_gss3:*
85: gb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: gb_gss5:*
91: gb_gss6:*
92: gb_gss7:*
93: gb_gss8:*
94: gb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: gb_gss10:*
103: gb_gss11:*
104: em_gss12:*
105: gb_gss12:*
106: gb_gss13:*
107: gb_gss14:*
108: gb_gss15:*
109: gb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	% Query

	No.	Score	Match	Length	DB	ID	Description
	1	96.4	0.4	106	37	AA703692	AA703692 ag81a10.r
	2	94.8	0.4	109	30	AA243009	AA243009 zr25h02.s
	3	92.4	0.4	103	38	AA807640	AA807640 nx08b05.s
	4	92.6	0.4	108	84	B65160	B65160 CIT-HSP-201
	5	92.4	0.4	110	39	AA897366	AA897366 am06h02.s
	6	91.4	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
c	7	91	0.4	108	84	B65160	B65160 CIT-HSP-201
	8	89.8	0.4	101	39	AA835205	AA835205 ak64h01.s
	9	89.8	0.4	110	30	AA244245	AA244245 nc07a04.s
	10	88.6	0.4	103	30	AA228795	AA228795 nc14e07.s
c	11	88.4	0.4	106	105	AQ264176	AQ264176 CITBI-E1-
c	12	87.8	0.4	103	84	B48914	B48914 RPCI11-4A12
	13	87.8	0.4	107	35	AA565533	AA565533 nk42b11.s
c	14	86.6	0.4	103	108	AQ535244	AQ535244 RPCI-11-3
	15	86.8	0.4	103	108	AQ582186	AQ582186 RPCI-11-4
c	16	87	0.4	103	108	AQ584425	AQ584425 RPCI-11-4
	17	86.8	0.4	106	38	AA812141	AA812141 ob48h02.s
	18	86.8	0.4	106	94	AQ062963	AQ062963 CIT-HSP-2
	19	86.4	0.4	105	30	AA218889	AA218889 zq15d04.s
c	20	86.4	0.4	107	39	AA828124	AA828124 od71a07.s
c	21	86.6	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
c	22	86.6	0.4	110	30	AA244245	AA244245 nc07a04.s
	23	85.6	0.4	102	30	AA226656	AA226656 nc19f09.s
	24	85.8	0.4	105	28	AA078003	AA078003 7H12D08 C
c	25	85.8	0.4	105	105	AQ282107	AQ282107 RPCI11-94
c	26	85.2	0.4	109	22	H11143	H11143 ym09c06.r1
	27	84.8	0.4	104	29	AA129957	AA129957 zn86h04.r
c	28	84.8	0.4	104	29	AA129957	AA129957 zn86h04.r
	29	84.8	0.4	107	33	AA385808	AA385808 EST99495
	30	85	0.4	109	105	AQ265749	AQ265749 CITBI-E1-
	31	84.6	0.4	107	103	AQ240182	AQ240182 CIT-HSP-2
	32	83.8	0.4	108	94	AQ014433	AQ014433 CIT-HSP-2
c	33	83.4	0.4	101	39	AA835205	AA835205 ak64h01.s
	34	83.4	0.4	102	94	AQ004934	AQ004934 CIT-HSP-2
	35	83.6	0.4	109	24	N25299	N25299 yw52c09.s1
c	36	83.4	0.4	109	84	B17434	B17434 345K2.TVB C
c	37	83	0.4	101	33	AA381369	AA381369 EST94442
c	38	83.2	0.4	104	108	AQ544583	AQ544583 CITBI-E1-
	39	83.2	0.4	107	24	N23686	N23686 yw46a02.s1
	40	83.4	0.4	110	94	AQ003188	AQ003188 RPCI11-1D
c	41	82.6	0.4	106	108	AQ544957	AQ544957 CITBI-E1-
	42	82.8	0.4	110	29	AA177157	AA177157 nc02g07.s
	43	82.4	0.4	101	94	AQ076649	AQ076649 CIT-HSP-2
	44	82	0.4	98	24	H67549	H67549 yu68f10.s1
	45	82.2	0.4	104	105	AQ268072	AQ268072 RPCI11-73

ALIGNMENTS

```
RESULT   1
AA703692
LOCUS       AA703692      106 bp    mRNA            EST       24-DEC-1997
DEFINITION  ag81a10.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            IMAGE:1140858 5' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA703692
VERSION     AA703692.1  GI:2713610
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1397630.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 53.
FEATURES             Location/Qualifiers
     source          1. .106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1140858"
                     /clone_lib="Stratagene hNT neuron (#937233)"
                     /dev_stage="hNT neurons"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Differentiated, post mitotic hNT neurons. Average insert
                     size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
                     GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     29 c     29 g     29 t
ORIGIN


  Query Match             0.4%;  Score 96.4;  DB 37;  Length 106;
  Best Local Similarity   94.3%;  Pred. No. 0.066;
  Matches  100;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy  10431 TTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 10490
          ||||||||||||||| ||||||||| |||||||||||||||||||||||||||||||||||
Db      1 TTTTTAGTAGAGACGAGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 60

Qy  10491 TGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGGTGTG 10536
          ||||| |||| ||| ||||||||||||||||||||||||||| |||
Db     61 TGATCTGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG 106


RESULT   2
AA243009
LOCUS       AA243009      109 bp    mRNA            EST       11-MAR-1998
DEFINITION  zr25h02.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
            cDNA clone IMAGE:664467 3' similar to contains Alu repetitive
            element;contains element LTR1 repetitive element ;, mRNA sequence.
ACCESSION   AA243009
VERSION     AA243009.1  GI:1873869
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Dec 3, 1996 this sequence version replaced gi:1126869.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1127    Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 102.
```

```
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="GDB:5426481"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:664467"
                     /clone_lib="Stratagene NT2 neuronal precursor 937230"
                     /tissue_type="neuroepithelial cells"
                     /dev_stage="Ntera-2 neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: brain; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. Uninduced, exponentially growing neuroepithelial
                     cells (Ntera-2/c1.D1). Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     30 c     30 g     30 t
ORIGIN


  Query Match             0.4%;  Score 94.8;  DB 30;  Length 109;
  Best Local Similarity   93.4%;  Pred. No. 0.098;
  Matches   99;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy  10431 TTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 10490
          ||||||||||||||||||||||| |||||||||||||||||  |||||||  ||||||
Db      4 TTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTTGATCTCCCTACCTCG 63

Qy  10491 TGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGGTGTG 10536
          |||||||||||||  |||||||||||||||||||||||||| |||
Db     64 TGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 109


RESULT   3
AA807640
LOCUS       AA807640      103 bp    mRNA            EST       05-MAR-1998
DEFINITION  nx08b05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA807640
VERSION     AA807640.1  GI:2877108
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2151346.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 774   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1255473"
                     /clone_lib="NCI_CGAP_GC3"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is not normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       19 a     27 c     30 g     27 t
ORIGIN


  Query Match             0.4%;  Score 92.4;  DB 38;  Length 103;
  Best Local Similarity   94.1%;  Pred. No. 0.18;
  Matches   96;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy  10436 AGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATC 10495
          |||||||| |||||||||| |||||||||||||||||||||||||||||||| ||||||
Db      2 AGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTTGTGATC 61

Qy  10496 CGCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGGTGTGA 10537
          ||| |||||  |||||||||||||||||||||||||||||||
Db     62 CGCTCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGTGTGA 103


RESULT   4
B65160
LOCUS       B65160        108 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
            genomic survey sequence.
ACCESSION   B65160
VERSION     B65160.1  GI:2639138
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2017G2.TFB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7043860"
                     /db_xref="taxon:9606"
                     /clone="2017G2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       26 a     27 c     34 g     21 t
ORIGIN
```

```
  Query Match             0.4%;  Score 92.6;  DB 84;  Length 108;
  Best Local Similarity   91.6%;  Pred. No. 0.17;
  Matches   98;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy   4994 TGGTGGCAGGCACCTGTAATCCCAGCTATTCAGGAGGCTGAGACAGGAGAATCGCTTGAA 5053
          |||||||| || |||||||||||||||| |||||||||||||| |||||||||| ||||||
Db      1 TGGTGGCATGCGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCTCTTGAA 60

Qy   5054 CCCAGGAGGTGGAGGTTGCATTGAGCCAAGATCATGCCACTGCACTC 5100
          ||| |||||||||||||| |||||||||||||| ||||||||| |
Db     61 CCCGGGAGGTGGAGGTTGCAGTGAGCCAAGATCATACCACTGCACAC 107


RESULT   5
AA897366
LOCUS       AA897366      110 bp    mRNA            EST       04-JAN-1999
DEFINITION  am06h02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1466067 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   AA897366
VERSION     AA897366.1  GI:3033986
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 19, 1998 this sequence version replaced gi:2150764.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 834   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 63.
FEATURES             Location/Qualifiers
     source          1. .110
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1466067"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT       22 a     27 c     29 g     32 t
ORIGIN


  Query Match             0.4%;  Score 92.4;  DB 39;  Length 110;
  Best Local Similarity   90.0%;  Pred. No. 0.18;
  Matches   99;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  10427 GTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGAC 10486
          |||||||||||||||| |||||||||| ||||| |||||||||||||| |||||  |||
Db      1 GTATTTTTTAGTAGAGATGGGGTTTCACCGTGTTAACCAGGATGGTCTCAATCTCTGGAC 60

Qy  10487 CTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGATTACAGGTGTG 10536
          ||| |||||||||||||  |||||||||||||||||||||| ||| |||
Db     61 CTCATGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTATAGGCGTG 110


RESULT   6
AQ028426
LOCUS       AQ028426      109 bp    DNA             GSS       30-JUN-1998
DEFINITION  CIT-HSP-2313G15.TF CIT-HSP Homo sapiens genomic clone 2313G15,
            genomic survey sequence.
ACCESSION   AQ028426
VERSION     AQ028426.1  GI:3268648
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 109)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building (1998)
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2313G15"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       19 a     36 c     25 g     29 t
ORIGIN


  Query Match             0.4%;  Score 91.4;  DB 94;  Length 109;
  Best Local Similarity   89.9%;  Pred. No. 0.23;
  Matches   98;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  17725 TTTTTTTTTGAGACGGAGTCTCGTTCTGTCACCCAGGCTGGAGTGCAGTGGCACAATCTC 17784
          || |||| ||||||||| |||| ||||||||||||||||||||||||||||||||||| |||
Db      1 TTGTTTTCTGAGACGGACTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCACAGTCTG 60

Qy  17785 GGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCTATTCTCCTGCCTCA 17833
          |||||||||||||||| ||||| |||||||||| ||||||||||||||
Db     61 AGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCA 109


RESULT   7
B65160/c
LOCUS       B65160        108 bp    DNA             GSS       21-JUN-1998
DEFINITION  CIT-HSP-2017G2.TRB CIT-HSP Homo sapiens genomic clone 2017G2,
            genomic survey sequence.
ACCESSION   B65160
VERSION     B65160.1  GI:2639138
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
            Simon,M. and Venter,J.C.
  TITLE     Use of a random BAC End Sequence Database for Sequence-Ready Map
            Building
  JOURNAL   Unpublished (1997)
COMMENT     Other_GSSs: CIT-HSP-2017G2.TFB
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Homo sapiens"
                     /db_xref="GDB:7043860"
                     /db_xref="taxon:9606"
                     /clone="2017G2"
                     /clone_lib="CIT-HSP"
                     /sex="Male"
                     /cell_type="Sperm"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII"
BASE COUNT       26 a     27 c     34 g     21 t
ORIGIN


  Query Match             0.4%;  Score 91;  DB 84;  Length 108;
  Best Local Similarity   90.7%;  Pred. No. 0.26;
  Matches   97;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  18250 GAGTGCAGTGGCATGATCTCGGCTCACAGCAACCTCCACCTCCTGGGTTCAAGTGATTCT 18309
          | ||||||||| ||||||| ||||||| ||||||||||||||| ||||||||| ||||||
Db    107 GTGTGCAGTGGTATGATCTTGGCTCACTGCAACCTCCACCTCCCGGGTTCAAGAGATTCT 48

Qy  18310 CCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGACGTGTGTCACCA 18356
          |||||||||||||||||||||||||||||||||| ||  || |||||
Db     47 CCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCATGCCACCA 1


RESULT   8
AA835205
LOCUS       AA835205      101 bp    mRNA            EST       23-FEB-1998
DEFINITION  ak64h01.s1 Barstead pancreas HPLRB1 Homo sapiens cDNA clone
            IMAGE:1412689 3' similar to contains Alu repetitive
            element;contains element KER repetitive element ;, mRNA sequence.
ACCESSION   AA835205
VERSION     AA835205.1  GI:2908933
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 101)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
            Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     On Nov 29, 1993 this sequence version replaced gi:636191.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .101
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1412689"
                     /clone_lib="Barstead pancreas HPLRB1"
                     /sex="female"
                     /dev_stage="adult, 34 years"
                     /lab_host="DH10B"
                     /note="Organ: pancreas; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: EcoRI; Site_2: NotI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead. "
BASE COUNT       14 a     36 c     27 g     24 t
ORIGIN


  Query Match             0.4%;  Score 89.8;  DB 39;  Length 101;
  Best Local Similarity   93.1%;  Pred. No. 0.36;
  Matches   94;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy  11783 TGAGACGGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACT 11842
          |||||||||||||||||||| |||||||||||||||||||  |||||||||||||
Db      1 TGAGACGGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCTTGATCTCGGCTCACT 60

Qy  11843 GCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCA 11883
          |||| |||||||||||||||| |  ||||||||||||||
Db     61 GCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCA 101


RESULT   9
AA244245
LOCUS       AA244245      110 bp    mRNA            EST       20-AUG-1997
DEFINITION  nc07a04.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1007406
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA244245
VERSION     AA244245.1  GI:1875104
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 110)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Jan 24, 1995 this sequence version replaced gi:634306.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 90.
FEATURES             Location/Qualifiers
```

```
    source          1. .110
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1007406"
                    /clone_lib="NCI_CGAP_Pr1"
                    /sex="Male"
                    /dev_stage="45 years old"
                    /lab_host="DH10B"
                    /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                    strand cDNA was primed with oligo(dT)17 on 50 ng of
                    DNAse-treated, total cellular RNA obtained from
                    5,000-10,000 microdissected, histologically normal
                    prostate epithelial cells. Double-stranded cDNA was
                    ligated to EcoRI adaptors, 5 cycles of PCR applied to the
                    cDNA with an adaptor-specific primer, and the resulting
                    PCR product subcloned into pAMP10 by the UDG-cloning
                    method (Life Technologies).  Average insert size is 600
                    bp. NOTE: Not directionally cloned.  This library was
                    constructed by David Krizman."
BASE COUNT       17 a     26 c     28 g     38 t      1 others
ORIGIN


  Query Match             0.4%;  Score 89.8;  DB 30;  Length 110;
  Best Local Similarity   88.2%;  Pred. No. 0.35;
  Matches   97;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy  17724 TTTTTTTTTTGAGACGGAGTCTCGTTCTGTCACCCAGGCTGGAGTGCAGTGGCACAATCT 17783
          |||||||||||||| ||||||| | |||||  |||||||||||||||||||||| | |||
Db      1 TTTTTTTTTTGAGATGGAGTCTTGATCTGTTGCCCAGGCTGGAGTGCAGTGGCAGANTCT 60

Qy  17784 CGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCTATTCTCCTGCCTCA 17833
          |||||||||||||||| |||||| |||||||||  ||||| ||||||||
Db     61 TGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGAGATTCTTCTGCCTCA 110


RESULT  10
AA228795
LOCUS       AA228795      103 bp    mRNA            EST       20-AUG-1997
DEFINITION  nc14e07.s1 NCI_CGAP_Pr1 Homo sapiens cDNA clone IMAGE:1008132
            similar to contains Alu repetitive element;contains element MER28
            repetitive element ;, mRNA sequence.
ACCESSION   AA228795
VERSION     AA228795.1  GI:1851455
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1394473.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 81.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1008132"
                     /clone_lib="NCI_CGAP_Pr1"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from
                     5,000-10,000 microdissected, histologically normal
                     prostate epithelial cells. Double-stranded cDNA was
                     ligated to EcoRI adaptors, 5 cycles of PCR applied to the
                     cDNA with an adaptor-specific primer, and the resulting
                     PCR product subcloned into pAMP10 by the UDG-cloning
                     method (Life Technologies).  Average insert size is 600
                     bp. NOTE: Not directionally cloned.  This library was
                     constructed by David Krizman."
BASE COUNT       14 a     30 c     25 g     34 t
ORIGIN


  Query Match             0.4%;  Score 88.6;  DB 30;  Length 103;
  Best Local Similarity   91.3%;  Pred. No. 0.48;
  Matches   94;  Conservative    0;  Mismatches   9;  Indels    0;  Gaps    0;

Qy   1519 TTTTTTTTTTTTGAGACGGTGTCTCACTCTTTCGCCCAGGCCGGACTGCAGTAGCGCTAT 1578
           |||||||||||||||| |||||||||||| |||||||||| ||| |||| |||||| ||
Db      1 TTTTTTTTTTTTGAGATGGTGTCTCACTCTGTCGCCCAGGCTGGAGTGCACTAGCGCAAT 60

Qy   1579 CTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTTT 1621
           || |||||||||||||||||||||||||||||||||||| || |
Db     61 CTTGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCGTTAT 103


RESULT  11
AQ264176/c
LOCUS       AQ264176      106 bp    DNA             GSS       27-OCT-1998
DEFINITION  CITBI-E1-2509A2.TF CITBI-E1 Homo sapiens genomic clone 2509A2,
            genomic survey sequence.
ACCESSION   AQ264176
VERSION     AQ264176.1  GI:3792743
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 106)
  AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
  TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1998)
COMMENT     Other_GSSs: CITBI-E1-2509A2.TR
            Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .106
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2509A2"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
```

```
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
                     CalTech Human BAC Library D"
BASE COUNT       25 a     30 c     34 g     17 t
ORIGIN


  Query Match             0.4%;  Score 88.4;  DB 105;  Length 106;
  Best Local Similarity   89.6%;  Pred. No. 0.5;
  Matches   95;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy  10446 GGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTG 10505
          ||||||||||||||||||||||| ||||| |||||||||||||||||||| ||| |||
Db    106 GGGTTTCACCATGTTAGCCAGGACGGTCTTGATCTCCTGACCTCGTGATCCACCCGCCTC 47

Qy  10506 AGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGG 10551
          | |||||||||||||||||||||||| ||||| | | ||||||||
Db     46 GGTCTCCCAAAGTGCTGGGATTACAGGCGTGAGACTCTGCGCCCGG 1


RESULT  12
B48914/c
LOCUS       B48914         103 bp    DNA             GSS       08-APR-1999
DEFINITION  RPCI11-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
            genomic survey sequence.
ACCESSION   B48914
VERSION     B48914.1  GI:2601151
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences for Sequence-Ready Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7501163"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-4A12"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       30 a     28 c     30 g     15 t
ORIGIN


  Query Match             0.4%;  Score 87.8;  DB 84;  Length 103;
  Best Local Similarity   92.9%;  Pred. No. 0.59;
  Matches   92;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy  10431 TTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCG 10490
          ||||||||||||||||||||||| | |||||| |||||||||||||||||||||||||
Db     99 TTTTTAGTAGAGACGGGGTTTCACCGTTTTAGCCGGGATGGTCTCGATCTCCTGACCTCG 40

Qy  10491 TGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGATTAC 10529
          |||||||||| ||| ||||||||||||||||||| ||||
Db     39 TGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGCTTAC 1


RESULT  13
AA565533
LOCUS       AA565533       107 bp    mRNA            EST       08-SEP-1997
DEFINITION  nk42b11.s1 NCI_CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
            similar to contains Alu repetitive element;, mRNA sequence.
ACCESSION   AA565533
VERSION     AA565533.1  GI:2337172
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 107)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1393355.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Insert Length: 1661   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .107
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1016157"
                     /clone_lib="NCI_CGAP_GC2"
                     /tissue_type="germ cell tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. Bulk
                     germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCACGAG 3'
                     3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                     Average insert size:  1.2 kb."
BASE COUNT       22 a     34 c     26 g     25 t
ORIGIN


  Query Match             0.4%;  Score 87.8;  DB 35;  Length 107;
  Best Local Similarity   88.8%;  Pred. No. 0.58;
  Matches   95;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  11812 CTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGAT 11871
          |||||||||||||||| | ||||| |||||||||||| ||| ||||||| ||||||||||||
Db      1 CTGGAGTGCAGTGGCTCAATCTCAGCTCACTGCAAGCTCTGCCTCCCAGGTTCAAGTGAT 60

Qy  11872 TCTCCTGCCTCAGACTCCCGAGTAGCTGGGATTACAGTCATGCACCA 11918
          |||| |||||||| |||| ||||||||||||||||| ||  |||||
Db     61 TCTCGTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACACACCA 107


RESULT  14
AQ535244/c
```

```
LOCUS       AQ535244      103 bp    DNA             GSS       18-MAY-1999
DEFINITION  RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
            RPCI-11-317H22, genomic survey sequence.
ACCESSION   AQ535244
VERSION     AQ535244.1  GI:4846934
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7621533"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-317H22"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       31 a     27 c     27 g     18 t
ORIGIN
```

```
  Query Match             0.4%;  Score 86.6;  DB 108;  Length 103;
  Best Local Similarity   91.1%;  Pred. No. 0.8;
  Matches   92;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy  14687 TTTTTGTATTTCTGGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCC 14746
          |||||||||||  | | ||||||||||||||||||||||||||||||||||||||||||
Db    103 TTTTTGTATTTATAGCAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCC 44

Qy  14747 TGACCTCAAGTGATCCACAATCCTTGGCCTCCCAAAGTGCT 14787
          |||||||||||||||  |    ||||||||||||||||||
Db     43 TGACCTCAAGTGATCTGCCCGTCTTGGCCTCCCAAAGTGCT 3
```

```
RESULT  15
AQ582186
LOCUS       AQ582186      103 bp    DNA             GSS       07-JUN-1999
DEFINITION  RPCI-11-451A15.TJ RPCI-11 Homo sapiens genomic clone
            RPCI-11-451A15, genomic survey sequence.
ACCESSION   AQ582186
VERSION     AQ582186.1  GI:5009296
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
            Venter,J.C.
  TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
  JOURNAL   Unpublished (1997)
COMMENT     On Feb 19, 1999 this sequence version replaced gi:4146076.
            Other_GSSs: RPCI-11-451A15.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Homo sapiens"
                     /db_xref="GDB:7672814"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-451A15"
                     /clone_lib="RPCI-11"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPCI11 Human Male BAC Library"
BASE COUNT       19 a     36 c     25 g     22 t      1 others
ORIGIN
```

```
  Query Match             0.4%;  Score 86.8;  DB 108;  Length 103;
  Best Local Similarity   91.9%;  Pred. No. 0.76;
  Matches   91;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy  17764 GGAGTGCAGTGGCACAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCTATTC 17823
          ||||||| |||||||||||||||||||||||||||||| ||||||| | ||||||| |||
Db      1 GGAGTGCNGTGGCACAATCTCGGCTCACTGCAACCTCTGCCTCCCAGATTCAAGCGATTT 60

Qy  17824 TCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGC 17862
          ||||||||||||||| |||||||||||||||| |||
Db     61 TCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGACGC 99
```

```
Search completed: June 17, 2000, 13:46:22
Job time: 263349 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:50:33 ; Search time 286.25 Seconds
(without alignments)
10005.207 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
Perfect score: 22034
Sequence: 1 TTCGGAGTGATTTTGGTCAT.....TGTGTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	76.2	0.3	105	4	US-08-481-658B-65	Sequence 65, Appl
	2	76.2	0.3	105	4	US-08-477-504A-65	Sequence 65, Appl
	3	76.2	0.3	105	4	US-08-486-756A-65	Sequence 65, Appl
	4	76.2	0.3	105	4	US-08-485-862B-65	Sequence 65, Appl
	5	76.2	0.3	105	5	US-08-787-739-65	Sequence 65, Appl
	6	67	0.3	84	3	US-08-454-557C-91	Sequence 91, Appl
	7	67	0.3	84	4	US-08-340-426D-91	Sequence 91, Appl
	8	67	0.3	84	4	US-08-450-673C-91	Sequence 91, Appl
	9	67	0.3	84	6	PCT-US95-17111A-91	Sequence 91, Appl
c	10	60.6	0.3	84	3	US-08-454-557C-91	Sequence 91, Appl
c	11	60.6	0.3	84	4	US-08-340-426D-91	Sequence 91, Appl
c	12	60.6	0.3	84	4	US-08-450-673C-91	Sequence 91, Appl
c	13	60.6	0.3	84	6	PCT-US95-17111A-91	Sequence 91, Appl
	14	58.8	0.3	78	3	US-08-454-557C-70	Sequence 70, Appl
	15	58.8	0.3	78	4	US-08-340-426D-70	Sequence 70, Appl
	16	58.8	0.3	78	4	US-08-450-673C-70	Sequence 70, Appl
	17	58.8	0.3	78	6	PCT-US95-17111A-70	Sequence 70, Appl
	18	57.6	0.3	85	3	US-08-454-557C-92	Sequence 92, Appl
	19	57.6	0.3	85	4	US-08-340-426D-92	Sequence 92, Appl
	20	57.6	0.3	85	4	US-08-450-673C-92	Sequence 92, Appl
	21	57.6	0.3	85	6	PCT-US95-17111A-92	Sequence 92, Appl
c	22	54.8	0.2	105	4	US-08-481-658B-65	Sequence 65, Appl
c	23	54.8	0.2	105	4	US-08-477-504A-65	Sequence 65, Appl
c	24	54.8	0.2	105	4	US-08-486-756A-65	Sequence 65, Appl
c	25	54.8	0.2	105	4	US-08-485-862B-65	Sequence 65, Appl
c	26	54.8	0.2	105	5	US-08-787-739-65	Sequence 65, Appl
c	27	54.4	0.2	106	3	US-08-332-766A-36	Sequence 36, Appl
	28	54	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
	29	54	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
	30	54	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
	31	54	0.2	76	6	PCT-US95-17111A-69	Sequence 69, Appl
	32	53.6	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
	33	53.6	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
	34	53.6	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
	35	53.6	0.2	60	6	PCT-US95-17111A-57	Sequence 57, Appl
c	36	50.4	0.2	60	5	US-08-545-860D-71	Sequence 71, Appl
c	37	50.4	0.2	60	6	PCT-US94-04496-71	Sequence 71, Appl
c	38	50	0.2	94	5	US-08-750-064-6	Sequence 6, Appli
	39	47.6	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
	40	47.6	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
	41	47.6	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
	42	47.6	0.2	60	6	PCT-US95-17111A-60	Sequence 60, Appl
	43	46.2	0.2	59	3	US-08-454-557C-65	Sequence 65, Appl
	44	46.2	0.2	59	4	US-08-340-426D-65	Sequence 65, Appl
	45	46.2	0.2	59	4	US-08-450-673C-65	Sequence 65, Appl

ALIGNMENTS

```
RESULT   1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,658B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3E
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-481-658B-65


  Query Match              0.3%;  Score 76.2;  DB 4;  Length 105;
  Best Local Similarity   82.9%;  Pred. No. 5.6e-07;
```

```
  Matches   87;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  10422 TTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10481
          |||||    | ||||||||||| |||||||||||  || ||||||  || ||||  | |||
Db      1 TTTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy  10482 CTGACCTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGAT 10526
          |||||||  |||||||  ||  |||  ||||||||||||||||||||
Db     61 CTGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   2
US-08-477-504A-65
; Sequence 65, Application US/08477504A
; Patent No. 5972353
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,504A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3D
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-477-504A-65


  Query Match             0.3%;  Score 76.2;  DB 4;  Length 105;
  Best Local Similarity   82.9%;  Pred. No. 5.6e-07;
  Matches   87;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  10422 TTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10481
          |||||    | ||||||||||| |||||||||||  || ||||||  || ||||  | |||
Db      1 TTTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy  10482 CTGACCTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGAT 10526
          |||||||  |||||||  ||  |||  ||||||||||||||||||||
Db     61 CTGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   3
US-08-486-756A-65
; Sequence 65, Application US/08486756A
; Patent No. 5981711
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
;      CITY:  Tiburon
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94920
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,756A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/260,190
;      FILING DATE:  15-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lauder, Leona L.
;      REGISTRATION NUMBER:  30,863
;      REFERENCE/DOCKET NUMBER:  D-0021.3C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-435-2034
;      TELEFAX:  415-435-0727
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-486-756A-65


  Query Match             0.3%;  Score 76.2;  DB 4;  Length 105;
  Best Local Similarity   82.9%;  Pred. No. 5.6e-07;
  Matches   87;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  10422 TTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10481
          |||||    | ||||||||||| |||||||||||  || ||||||  || ||||  | |||
Db      1 TTTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy  10482 CTGACCTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGAT 10526
          |||||||  |||||||  ||  |||  ||||||||||||||||||||
Db     61 CTGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   4
US-08-485-862B-65
; Sequence 65, Application US/08485862B
; Patent No. 5989838
;  GENERAL INFORMATION:
;    APPLICANT:  Zavada, Jan
;    APPLICANT:  Pastorekova, Silvia
;    APPLICANT:  Pastorek, Jaromir
;    TITLE OF INVENTION:  MN Gene and Protein
;    NUMBER OF SEQUENCES:  86
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leona L. Lauder
;      STREET:  6 Mariposa Court
```

```
;       CITY:  Tiburon
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94920
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,862B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;       APPLICATION NUMBER:  US 08/260,190
;       FILING DATE:  15-JUN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.3D
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-435-2034
;       TELEFAX:  415-435-0727
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-485-862B-65


  Query Match             0.3%;  Score 76.2;  DB 4;  Length 105;
  Best Local Similarity   82.9%;  Pred. No. 5.6e-07;
  Matches   87;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  10422 TTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10481
          |||||    | |||||||||||| |||||||||||| || |||||| || |||| | |||
Db      1 TTTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy  10482 CTGACCTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGAT 10526
          ||||||| ||||||| ||  |||  ||||||||||||||||||||
Db     61 CTGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
;  GENERAL INFORMATION:
;     APPLICANT:  Zavada, Jan
;     APPLICANT:  Pastorekova, Silvia
;     APPLICANT:  Pastorek, Jaromir
;     TITLE OF INVENTION:  MN Gene and Protein
;     NUMBER OF SEQUENCES:  96
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Leona L. Lauder
;       STREET:  369 Pine Street, Suite 610
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/787,739
;       FILING DATE:  24-JAN-1997
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,049
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/486,756
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/477,504
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/481,658
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,862
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,863
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/487,077
;       FILING DATE:  07-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Lauder, Leona L.
;       REGISTRATION NUMBER:  30,863
;       REFERENCE/DOCKET NUMBER:  D-0021.4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-981-2034
;       TELEFAX:  415-981-0332
;   INFORMATION FOR SEQ ID NO:  65:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  105 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-787-739-65


  Query Match             0.3%;  Score 76.2;  DB 5;  Length 105;
  Best Local Similarity   82.9%;  Pred. No. 5.6e-07;
  Matches   87;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy  10422 TTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTC 10481
          |||||    | |||||||||||| |||||||||||| || |||||| || |||| | |||
Db      1 TTTTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy  10482 CTGACCTCGTGATCCGCCCACCTGAGCCTCCCAAAGTGCTGGGAT 10526
          ||||||| ||||||| ||  |||  ||||||||||||||||||||
Db     61 CTGACCTTGTGATCCACCAGCCTCGGCCTCCCAAAGTGCTGGGAT 105


RESULT   6
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;     APPLICANT:  de la Monte, Suzanne
;     APPLICANT:  Wands, Jack R.
;     TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;     TITLE OF INVENTION:  of Alzheimer's Disease
;     NUMBER OF SEQUENCES:  121
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;       STREET:  1100 New York Avenue, Suite 600
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/454,557C
;       FILING DATE:  30-MAY-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ludwig, Steven R.
;       REGISTRATION NUMBER:  36,203
;       REFERENCE/DOCKET NUMBER:  0609.3840003
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  91:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  84 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
US-08-454-557C-91


  Query Match             0.3%;  Score 67;  DB 3;  Length 84;
  Best Local Similarity   88.0%;  Pred. No. 3.6e-05;
  Matches   73;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  10454 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAGCCTCCC 10513
          |||||||   |||| |||| |||| ||||||||||||||||||||||| ||| ||||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  10514 AAAGTGCTGGGATTACAGGTGTG 10536
          |||||||||||||||| | |||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT   7
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
US-08-340-426D-91


  Query Match             0.3%;  Score 67;  DB 4;  Length 84;
  Best Local Similarity   88.0%;  Pred. No. 3.6e-05;
  Matches   73;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  10454 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAGCCTCCC 10513
          |||||||   |||| |||| |||| ||||||||||||||||||||||| ||| ||||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  10514 AAAGTGCTGGGATTACAGGTGTG 10536
          |||||||||||||||| | |||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT   8
US-08-450-673C-91
; Sequence 91, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-91


  Query Match             0.3%;  Score 67;  DB 4;  Length 84;
  Best Local Similarity   88.0%;  Pred. No. 3.6e-05;
  Matches   73;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  10454 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAGCCTCCC 10513
          |||||||   |||| |||| |||| ||||||||||||||||||||||| ||| ||||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  10514 AAAGTGCTGGGATTACAGGTGTG 10536
          |||||||||||||||| | |||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83
```

```
RESULT  9
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-91


  Query Match             0.3%;  Score 67;  DB 6;  Length 84;
  Best Local Similarity   88.0%;  Pred. No. 3.6e-05;
  Matches   73;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy  10454 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCACCTGAGCCTCCC 10513
          |||||||   |||| |||| |||| ||||||||||||||||||||||| ||| ||||||||
Db      1 CCATGTTCATCAGGCTGGTGTCGAACTCCTGACCTCGTGATCCGCCCGCCTCAGCCTCCC 60

Qy  10514 AAAGTGCTGGGATTACAGGTGTG 10536
          |||||||||||||||| | |||
Db     61 AAAGTGCTGGGATTACAAGCGTG 83


RESULT  10
US-08-454-557C-91/c
; Sequence 91, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,557C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-454-557C-91


  Query Match             0.3%;  Score 60.6;  DB 3;  Length 84;
  Best Local Similarity   83.1%;  Pred. No. 0.0007;
  Matches   69;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   4867 CATGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAATG 4926
          || || |||||| |||||||||||||||||| ||||| ||| ||||| |||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy   4927 CAAGACCAGCCTGACCAATATGG 4949
          | | ||||||||||  || ||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  11
US-08-340-426D-91/c
; Sequence 91, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
```

```
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-91


  Query Match             0.3%;  Score 60.6;  DB 4;  Length 84;
  Best Local Similarity   83.1%;  Pred. No. 0.0007;
  Matches   69;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   4867 CATGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAATG 4926
          || || |||||| |||||||||||||||||| ||||| ||| ||||| |||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy   4927 CAAGACCAGCCTGACCAATATGG 4949
          | | ||||||||||  || ||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  12
US-08-450-673C-91/c
; Sequence 91, Application US/08450673C
; Patent No. 5948888
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/450,673C
;      FILING DATE:  30-MAY-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840004
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-450-673C-91


  Query Match             0.3%;  Score 60.6;  DB 4;  Length 84;
  Best Local Similarity   83.1%;  Pred. No. 0.0007;
  Matches   69;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   4867 CATGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAATG 4926
          || || |||||| |||||||||||||||||| ||||| ||| ||||| |||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy   4927 CAAGACCAGCCTGACCAATATGG 4949
          | | ||||||||||  || ||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  13
PCT-US95-17111A-91/c
; Sequence 91, Application PC/TUS9517111A
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and
;    TITLE OF INVENTION:  Detection of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17111A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/340,426
;      FILING DATE:  14-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  91:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  84 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
PCT-US95-17111A-91


  Query Match             0.3%;  Score 60.6;  DB 6;  Length 84;
  Best Local Similarity   83.1%;  Pred. No. 0.0007;
  Matches   69;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   4867 CATGCCTGTAATTCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCATGAGGTCAGGAATG 4926
          || || |||||| |||||||||||||||||| ||||| ||| ||||| |||||||||| |
Db     83 CACGCTTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGTT 24

Qy   4927 CAAGACCAGCCTGACCAATATGG 4949
          | | ||||||||||  || ||||
Db     23 CGACACCAGCCTGATGAACATGG 1


RESULT  14
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
```

```
;       STREET:  1100 New York Avenue, Suite 600
;       CITY:  Washington
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20005-3934
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/454,557C
;       FILING DATE:  30-MAY-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ludwig, Steven R.
;       REGISTRATION NUMBER:  36,203
;       REFERENCE/DOCKET NUMBER:  0609.3840003
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202) 371-2600
;       TELEFAX:  (202) 371-2540
;   INFORMATION FOR SEQ ID NO:  70:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  78 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
US-08-454-557C-70


  Query Match             0.3%;  Score 58.8;  DB 3;  Length 78;
  Best Local Similarity   84.6%;  Pred. No. 0.0016;
  Matches   66;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  14672 ACCACGCCTGGCTAATTTTTGTATTTCTGGTAGAGACGGGGTTTCACCATGTTGGCCAGG 14731
          || |||||  |||||| ||||||||| | ||||||| |||||||| |||||||   ||||
Db      1 ACAACGCCCAGCTAATATTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTCATCAGG 60

Qy  14732 CTGGTCTCGAACTCCTGA 14749
          ||||| ||||||||||||
Db     61 CTGGTGTCGAACTCCTGA 78


RESULT  15
US-08-340-426D-70
; Sequence 70, Application US/08340426D
; Patent No. 5948634
;  GENERAL INFORMATION:
;    APPLICANT:  de la Monte, Suzanne
;    APPLICANT:  Wands, Jack R.
;    TITLE OF INVENTION:  Neural Thread Protein Gene Expression and Detection
;    TITLE OF INVENTION:  of Alzheimer's Disease
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/340,426D
;      FILING DATE:  14-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ludwig, Steven R.
;      REGISTRATION NUMBER:  36,203
;      REFERENCE/DOCKET NUMBER:  0609.3840002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  70:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  78 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-340-426D-70


  Query Match             0.3%;  Score 58.8;  DB 4;  Length 78;
  Best Local Similarity   84.6%;  Pred. No. 0.0016;
  Matches   66;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy  14672 ACCACGCCTGGCTAATTTTTGTATTTCTGGTAGAGACGGGGTTTCACCATGTTGGCCAGG 14731
          || |||||  |||||| ||||||||| | ||||||| |||||||| |||||||   ||||
Db      1 ACAACGCCCAGCTAATATTTGTATTTTTAGTAGAGATGGGGTTTCTCCATGTTCATCAGG 60

Qy  14732 CTGGTCTCGAACTCCTGA 14749
          ||||| ||||||||||||
Db     61 CTGGTGTCGAACTCCTGA 78


Search completed: June 17, 2000, 17:39:10
Job time: 276280 sec
```